

09/1189760

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:42:21 ; Search time 24 Seconds  
(without alignments)  
36.063 Million cell updates/sec

Title: SEQ:A  
Perfect score: 45  
Sequence: 1 vhhddlla 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 789

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	40.0	7	2 I46868	alpha-myosin heavy
2	17	37.8	9	2 S55696	phosphoenolpyruvat
3	16	35.6	7	2 PQ5663	membrane protein -
4	16	35.6	7	2 S68004	hucolin, 75K chain
5	16	35.6	8	2 PC4331	hypothetical prote
6	15	33.3	6	2 S78764	ribosomal protein
7	14	31.1	7	2 S20446	elastase - Pseudom
8	14	31.1	8	2 XGR0EU	urine glycopeptide
9	14	31.1	9	2 A12872	transaldolase (EC
10	14	31.1	9	2 S10920	venom protein HR-3
11	14	31.1	9	2 A60427	macrophage cyctox
12	13	28.9	8	2 A61328	trypsin (EC 3.4.21
13	13	28.9	9	2 PH0942	T-cell receptor be
14	12	26.7	4	2 I40697	bioitin A - Citroba
15	12	26.7	5	2 PT0601	T-cell receptor be
16	12	26.7	5	2 PT0679	T-cell receptor be
17	12	26.7	6	2 B35640	cerebellar degener
18	12	26.7	6	2 PT0533	T-cell receptor be
19	12	26.7	7	2 A34026	acetylcholinestera
20	12	26.7	7	2 B39040	calsequestrin, fas
21	12	26.7	7	2 PT0628	T-cell receptor be
22	12	26.7	7	2 PT0722	T-cell receptor be
23	12	26.7	7	2 PT0576	T-cell receptor be
24	12	26.7	8	2 PT0368	ig gamma chain C r
25	12	26.7	8	2 PN0043	phosphatidylethano
26	12	26.7	8	2 PT0557	T-cell receptor be
27	12	26.7	9	2 PW0002	chlorophyll a/b-bi
28	12	26.7	9	2 A60108	exotoxin A - Strep
29	12	26.7	9	2 S65913	pyrimidine synthe

30 12 26.7 9 2 PH0138 late G1-59 protein  
31 12 26.7 9 2 PT0562 T-cell receptor be  
32 12 26.7 9 2 B30572 T-cell receptor be  
33 11 24.4 5 2 C41225 copper resistance  
34 11 24.4 5 2 T10954 hypothetical prote  
35 11 24.4 6 2 T11779 phosphoglycerate t  
36 11 24.4 7 2 S25266 pLE protein - Bsc  
37 11 24.4 7 2 PT0246 ig heavy chain CRD  
38 11 24.4 8 2 S22428 chitin-binding pro  
39 11 24.4 8 2 B33099 158K exoantigen -  
40 11 24.4 8 2 S69165 ferredoxin a2 - Ja  
41 11 24.4 9 2 S66419 tetrameric protein  
42 11 24.4 9 2 A11497 transaldolase (EC  
43 11 24.4 9 2 PT0272 ig heavy chain CRD  
44 11 24.4 9 2 B39504 octamer-binding pr  
45 10 22.2 4 2 A48360 gamma subunit of p

ALIGNMENTS

RESULT 1  
I46868  
alpha-myosin heavy chain - rabbit (fragment)  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 14-Feb-1997 #sequence\_revision 14-Feb-1997 #text\_change 05-Nov-1999  
C:Accession: I46868  
R:Friedman, D.J.; Umeda, P.K.; Sinha, A.M.; Hsu, H.  
Proc. Natl. Acad. Sci. U.S.A. 81, 3044-3048, 1984  
A>Title: Characterization of genomic clones specifying rabbit alpha- and beta-ventricu  
A:Reference number: I46868; MUID:84221901; PMID:6328491  
A:Accession: I46868  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-7 <PRI>  
A:Cross-references: GB:K01698; NID:G165538; PIDN:AAA31415.1; PID:G165539

Query Match 40.0%; Score 18; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHDD 5  
DB 3 VHDE 6

RESULT 2  
S55696  
phosphoenolpyruvate carboxykinase - Trypanosoma brucei  
C:Species: Trypanosoma brucei  
C>Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-May-1999  
C:Accession: S55696  
R:Hunt, M.; Koehler, P.  
Biochim. Biophys. Acta 1249, 15-22, 1995  
A>Title: Purification and characterization of phosphoenolpyruvate carboxykinase from T  
A:Reference number: S55696; MUID:95284106; PMID:7766679  
A:Accession: S55696  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-9 <HUN>

Query Match 37.8%; Score 17; DB 2; Length 9;  
Best Local Similarity 33.3%; Pred. No. 2.8e+05;  
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLHDDL 6  
DB 3 ITHKXL 8

RESULT 3  
PQ0663  
membrane protein - porcine epidemic diarrhea virus (isolate Belgian CV777) (fragment)

```

C.Species: porcine epidemic diarrhea virus
C.Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 08-Oct-1999
C.Accession: PQ0663
C.Title: Hucclin, 75K chain - human (fragment)
A:Accession: PQ0663
A.Molecule type: mRNA
A.Residues: 1-7 <BR>
A.Cross-references: CB:Z14976; NID:g311650; PIDN:CAA78699.1; PID:G584283
A.Comment: This virus is coronavirus related to human coronavirus 229E.
C.Keywords: membrane protein

Query Match      35.6%; Score 16; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLH 3
   |||
Db 3 VLH 5
   |||

RESULT 4
S68004
nuclein, 75K chain - human (fragment)
C.Species: Homo sapiens (man)
C.Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C.Accession: S68004
R.Edgar, P.F.
PDB Lett. 375, 159-161, 1995
A>Title: Hucclin, a new corticosteroid-binding protein from human plasma with structural
A.Reference number: S68004; MUID:96087107; PMID:7498469
A.Accession: S68004
A>Status: preliminary
A.Molecule type: protein
A.Residues: 1-7 <EDG>

Query Match      35.6%; Score 16; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6
   |||
Db 4 DDL 6
   |||

RESULT 5
PC4131
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)
C.Species: Pseudomonas aeruginosa
C.Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C.Accession: PC4131
R.Kawasaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.
Gene 167, 87-91, 1995
A>Title: Sequencing and characterization of the downstream region of the genes encoding
y for biosynthesis of heme d1.
A.Reference number: Jc4552; MUID:96144254; PMID:8566817
A.Accession: PC4131
A>Status: preliminary
A.Molecule type: DNA
A.Residues: 1-8 <KAW>
A.Cross-references: DDBJ:250473; NID:g1217594
A.Note: this ORF is not annotated in GenBank entry PSENIRC, release 113.0

Query Match      35.6%; Score 16; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6
   |||
Db 2 DDL 4

```

---

```

RESULT 6
S78764
ribosomal protein XRP-S23, mitochondrial - bovine (fragment)
C.Species: Bos primigenius taurus (cattle)
C.Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C.Accession: S78764
R.Graack, H.R.
submitted to the Protein Sequence Database, July 1999
A:Reference number: S78760
A:Accession: S78764
A.Molecule type: protein
A.Residues: 1-6 <GRA>
C.Keywords: mitochondrion
F.I-6/Product: ribosomal protein XRP-S23 (fragment) #status experimental <MAT>

Query Match      33.3%; Score 15; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LHD 5
   |||
Db 2 LHD 4
   |||

RESULT 7
S20446
elastase - Pseudomonas aeruginosa
C.Species: Pseudomonas aeruginosa
C.Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 24-Jul-1997
C.Accession: S20446
R.Kessler, E.; Safran, M.; Peretz, M.; Burstein, Y.
PDB Lett. 299, 291-293, 1992
A>Title: Identification of cleavage sites involved in proteolytic processing of Pseudom
A.Reference number: S20446; MUID:92183956; PMID:1544509
A:Accession: S20446
A>Status: preliminary
A.Molecule type: protein
A.Residues: 1-7 <RES>

Query Match      31.1%; Score 14; DB 2; Length 7;
Best Local Similarity 50.3%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLJE 6
   |||
Db 3 DLJD 6
   |||

RESULT 8
XGHUEU
urine glycopeptide - human
C.Species: Homo sapiens (man)
C.Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 20-Jun-2000
C.Accession: A03188
R.Richte, C.J.; Weiss, J.B.
Biochem. J. 223, 25P, 1971
A>Title: Identification in urine of a low-molecular-weight polar glycopeptide containin
re has also been found (see FIR:XGHUE).
C.Superfamily: unassigned animal peptides
C.Keywords: glycoprotein
F.I-/Binding site: carbohydrate (Cys) (covalent) #status experimental

Query Match      31.1%; Score 14; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY      3 HD 4
      ||
Db      5 HD 6

RESULT 9
A12872
transaldolase (EC 2.2.1.2) I - yeast (Pichia jadinii) (fragment)
C:Species: Pichia jadinii, Candida utilis
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 30-Sep-1993
C:Accession: A12872
R:Sun, S.C.; Joris, L.; Tschäas, O.
Arch. Biochem. Biophys. 178, 69-78, 1977
A>Title: Purification and crystallization of transaldolase isozyme I and evidence for di
A:Reference number: A12872; MUID:77110646; PMID:556924
A:Accession: A12872
A:Molecule type: protein
A:Residues: 1-9 <SUN>
C:Keywords: transferase

Query Match      31.1%; Score 14; DB 2; Length 9;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 LHDLL 7
      ||
Db      3 IHCBTL 8

RESULT 10
S10920
venom protein HR-3 - oriental hornet (fragment)
C:Species: Vespa orientalis (oriental hornet)
C>Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 08-Dec-1995
C:Accession: S10920
R:Tsichibayev, M.U.; Akhmedova, N.U.; Kazakov, I.; Korreev, A.S.; Gagei'gans, A.I.
Biochemistry (N.Y.) 53, 183-190, 1988
A>Title: Low-molecular-weight peptides of venom of the giant hornet Vespa orientalis. S
A:Reference number: S06445
A:Accession: S10920
A:Molecule type: protein
A:Residues: 1-9 <TUI>
C:Keywords: venom

Query Match      31.1%; Score 14; DB 2; Length 9;
Best Local Similarity 28.6%; Pred. No. 2.8e+05;
Matches 2; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 LHDLL 8
      ||
Db      3 VHEFLVK 9

RESULT 11
A60427
macrophage cytotoxicity-inducing factor, 29K - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 18-Jun-1993
C:Accession: A60427
R:Jones, C.M.; Prince, C.A.; Williams, J.S.
Exp. Hematol. 19, 704-709, 1991
A>Title: Purification and amino acid analysis of a human macrophage cytotoxicity-inducin
A:Reference number: A60427; MUID:91372335; PMID:1909970
A:Accession: A60427
A:Molecule type: protein
A:Residues: 1-9 <JON>
A>Note: the sequence from the text on page 706 is inconsistent with that from page 708
C:Keywords: cytokine

Query Match      31.1%; Score 14; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 HD 4
      ||
Db      5 HD 6

RESULT 12
A61328
trypsin (EC 3.4.21.4) precursor - minke whale (fragment)
C:Species: Balenoptera acutorostrata (minke whale, lesser rorqual)
C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
C:Accession: A61328
R:Bricteux-Gregoire, S.; Schyns, R.; Florquin, M.; Emmens, M.; Weiling, G.W.; Beintema,
Biochim. Biophys. Acta 386, 244-255, 1975
A>Title: N-terminal amino acid sequence of trypsinogen from the lesser rorqual, Balae
C:Keywords: hydrolase; protein digestion; serine proteinase; zymogen

Query Match      28.9%; Score 13; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LHDD 5
      ||
Db      3 IDDD 6

RESULT 13
PH0942
T-cell receptor beta chain V-D-J region (clone 13) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C:Accession: PH0942
R:Gold, D.P.; Offer, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A>Title: Analysis of T cell receptor beta chains in Lewis rats with experimental aller
A:Reference number: PH0941; MUID:92078857; PMID:1836012
A:Accession: PH0942
A:Molecule type: mRNA
A:Residues: 1-9 <GOL>
A:Experimental source: complete Freund's adjuvant-immunized lymph node
A>Note: the authors translated the codon TGC for residue 2 as Ala
C:Keywords: T-cell receptor

Query Match      28.9%; Score 13; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 LJE 8
      ||
Db      5 LJE 7

RESULT 14
I40697
biotin A - Citrobacter freundii (fragment)
C:Species: Citrobacter freundii
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-1996
C:Accession: I40697
R:Shiuan, D.; Campbell, A.
Gene 67, 203-211, 1988
A>Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobac
A:Reference number: I40697; MUID:89006280; PMID:2971595
A:Accession: I40697
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4 <RES>
A:Cross-references: GB:M21922; NID:G144434
```

```

Query Match      26.7%; Score 12; DB 2; Length 4;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      4 DD 5
      ||
DB      3 DD 4

RESULT 15
PT0601
T-cell receptor beta chain V-D-J region (120-2K) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0601; PT0617; PT0694
R:Feeney, A.J.
J. Exp. Med. 174, 1-5-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0601
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <PEE>
A:Experimental source: newborn thymus, strain BALB/c, clone 120-2K
A:Accession: PT0617
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <PE3>
A:Experimental source: newborn thymus, strain BALB/c, 120-2CA
A:Accession: PT0694
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FE2>
A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-1H
C:Keywords: T-cell receptor

Query Match      26.7%; Score 12; DB 2; Length 5;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      4 DD 5
      ||
DB      4 DD 5

Search completed: October 20, 2003, 13:49:19
Job time : 24 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:32:06 ; Search time 13 Seconds

(without alignments)  
32.557 Million cell updates/sec

Title: SEQ1A  
Perfect score: 45  
Sequence: 1 vlnhdillea 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 251

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	16	35.6	9	1	FAR8_MACRS
2	14	31.1	6	1	TRPI_PSEPU
3	14	31.1	8	1	GLUR_HUMAN
4	14	31.1	9	1	TAL3_PICJA
5	12	26.7	5	1	BICA_CITFR
6	12	26.7	6	1	TMOP_SARBU
7	12	26.7	7	1	UF03_MOUSE
8	12	26.7	9	1	NSK1_SARBU
9	12	26.7	9	1	RE42_LITRU
10	12	26.7	9	1	UF02_MOUSE
11	12	26.7	9	1	UHAR_HUMAN
12	11	24.4	7	1	FAR2_ASCSJ
13	11	24.4	9	1	PSP_BOXMO
14	11	24.4	9	1	TAL3_PICJA
15	10	22.2	7	1	ALL7_CYPDO
16	10	22.2	7	1	HY7_PIG
17	10	22.2	7	1	PPH2_LYCES
18	10	22.2	7	1	UN06_PINPS
19	10	22.2	8	1	ANG2_BOTJA
20	10	22.2	8	1	P-IP_BRANA
21	10	22.2	8	1	UF05_MOUSE
22	10	22.2	8	1	KFI_BERAT
23	10	22.2	9	1	FIBB_LYCPA
24	10	22.2	9	1	LMIP_LOCM1
25	10	22.2	9	1	ULAH_HUMAN
26	9	20.0	5	1	E103_LITRU
27	9	20.0	5	1	E104_LITRU
28	9	20.0	5	1	TRM2_ECOLI
29	9	20.0	6	1	ASP2_LACSN
30	9	20.0	7	1	ALL2_CARMA
31	9	20.0	7	1	FAR1_HELTI
32	9	20.0	7	1	FAR8_CALVO
33	9	20.0	8	1	FAR8_CALVO

34	9	20.0	8	1	FUSS_FUSSO
35	9	20.0	9	1	ALC_CHLRE
36	9	20.0	9	1	FAR4_CALVO
37	9	20.0	9	1	FIBB_PAPHA
38	9	20.0	9	1	IPYR_RHOVI
39	9	20.0	9	1	PPH1_LYCES
40	9	20.0	9	1	ULAE_HUMAN
41	9	17.8	3	1	GSWM_HUMAN
42	8	17.8	3	1	THYL_PIG
43	8	17.8	4	1	DCML_PSECH
44	8	17.8	5	1	SIOP_CITFR
45	8	17.8	6	1	ACPH_RABIT

ALIGNMENTS

RESULT 1  
FAR8\_MACRS STANDARD; PRT; 9 AA.  
AC P83281.  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE FMRPamide-like neuropeptide FLP9 (VSHNNFLRF-amide).  
CS Macrobrachium rosenbergii (Giant fresh water prawn).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;  
OC Palaemonoidea; Palaemonidae; Macrobrachium.  
CX NCB1\_TaxID=79674;  
RN [1]  
RP SEQUENCE AND MASS SPECTROMETRY.  
RC TISSUE=Eyestalk;  
RX MEDLIN=21107394; PubMed=1179812;  
RA Sithigoringul P., Sarathongkum W., Longyant S., Panchan N.,  
RA Sithigoringul W., Petsom A.;  
RT "Three more novel FMRPamide-like neuropeptide sequences from the  
RT eyestalk of the giant freshwater prawn Macrobrachium rosenbergii.";  
RL Peptides 22:191-197(2001).  
CC -- MASS SPECTROMETRY: MW=1133.8; METHOD=MALDI.  
CC -- SIMILARITY: BELONGS TO THE FARP (FMRPAMIDE RELATED PEPTIDE)  
CC FAMILY.  
CC GO: GO:0007218; P:neuropeptide signaling pathway; IDA.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 9 AMIDATION.  
SQ SEQUENCE 9 AA; 1133 MW; 845A0729C4441F5 CRC64;

Query Match 35.6%; Score 16; DB 1; Length 9;  
Best Local Similarity 42.9%; Pred. No. 1.3e+05;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY : VJHDLJ 7  
DB : VSHNNFL 7

RESULT 2  
TRPI\_PSEPU STANDARD; PRT; 6 AA.  
ID TRPI\_PSEPU  
AC P36414;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE TrpBA operon transcriptional activator (fragment).  
GN TRPI.  
CS Pseudomonas putida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
CX NCB1\_TaxID=301;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PPG1 C:S;  
RX MEDLINE=89335826; PubMed=2503057;

```
RA Eberly L., Crawford I.P.;
RT "DNA sequence of the tryptophan synthase genes of Pseudomonas
RL putida";
RL Biochemie 71:521-531(1989).
CC !- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPBA GENES ENCODING
CC THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE
CC INDUCER (INDOLEGLYCEROL PHOSPHATE), TRPI BINDS UPSTREAM OF THE
CC TRAP OPERON, OVERLAPPING ITS OWN PROMOTER REGION.
CC !- SIMILARITY: BELONGS TO THE TYR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: X13299; CAA31660.1; .
CC InterPro: IPR000847; HTH LYSR.
CC PROSITE: PS00044; HTH LYSR FAMILY; PARTIAL.
CC Tryptophan biosynthesis; Transcription regulation; Activator;
CC DNA-binding.
CC NON TER 6 6
CC SEQUENCE 6 AA; 693 MW; 77672AA1EDD6F000 CRC64;
CC
CC Query Match 31.1%; Score 14; DB 1; Length 6;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+05;
CC Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 3 HD 4
CC 1 1
CC 3 HD 4
CC
CC Db
CC
CC RESULT 3
CC GLUR HUMAN STANDARD; PRT; 8 AA.
CC ID GLUR HUMAN STANDARD; PRT; 8 AA.
CC AC P02729;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE "Identification in urine of a low-molecular-weight highly polar
CC glycopeptide containing cysteine"-galactose.";
CC Biochem.J. 123:25P-25P(1971).
CC !- FUNCTION: THE IDENTITY OF THE GLYCOPROTEIN FROM WHICH THIS PEPTIDE
CC IS DERIVED IS UNKNOWN. NO PHYSIOLOGICAL FUNCTION HAS BEEN
CC ATTRIBUTED. AN ERYTHROCYTE MEMBRANE GLYCOPROTEIN HAVING A
CC SIMILAR STRUCTURE HAS ALSO BEEN FOUND.
CC
CC PIR: A03188; XGHUEU.
CC GO: G000576; C:extracellular; NAS.
CC Glycoprotein.
CC CARBOHYD 1 1 S-LINKED (GAL. .).
CC SEQUENCE 8 AA; 855 MW; C2D87AA1F5B1EB:E CRC64;
CC
CC Query Match 31.1%; Score 14; DB 1; Length 8;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+05;
CC Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 3 HD 4
CC 1 1
CC 5 HD 6
CC
CC Db
```

```
RESULT 4
ID TALI PICUA STANDARD; PRT; 9 AA.
AC P1740;
CC DT 01-AUG-1990 (Rel. 15, Created)
CC DT 01-AUG-1990 (Rel. 15, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Transaldolase I (EC 2.2.1.2) (Fragment).
CC OS Pichia jadinii (Yeast) (Candida utilis).
CC CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetaceae; Saccharomycetaceae; Pichia.
CC NCBI_TaxID=4903;
CC !-
CC MEDLINE=77110646; PubMed=556924;
CC Sun S.C., Joris L., Tselas O.;
CC "Purification of crystallization of transaldolase isozyme I and III in
CC evidence for different genetic origin of isozymes I and III in
CC Cardica utilis";
CC Arch. Biochem. Biophys. 178:69-78(1977).
CC !- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
CC !- CATALYTIC ACTIVITY: Sedheptulose 7-phosphate + D-glyceraldehyde
CC 3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.
CC !- PATHWAY: Pentose phosphate pathway; nonoxidative part.
CC !- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY. SUBFAMILY 1.
CC PIR: A12872; A12872.
CC InterPro: IPR001585; Transaldolase.
CC PROSITE: PS00958; TRANSALDOLASE_2; PARTIAL.
CC PROSITE: PS01054; TRANSALDOLASE_1; PARTIAL.
CC Cramer; Pentose shunt.
CC NON TER 1 1
CC NON TER 9 9
CC SEQUENCE 9 AA; 1008 MW; 274F31AF0EB1E058 CRC64;
CC
CC Query Match 31.1%; Score 14; DB 1; Length 9;
CC Best Local Similarity 33.3%; Pred. No. 1.3e+05;
CC Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 2 LHDDL 7
CC 1 1
CC 3 INCBTL 8
CC
CC Db
CC
CC RESULT 5
CC BIOA CITFR STANDARD; PRT; 5 AA.
CC ID BIOA CITFR STANDARD; PRT; 5 AA.
CC AC P13071;
CC DT 01-JAN-1990 (Rel. 13, Created)
CC DT 01-JAN-1990 (Rel. 13, Last sequence update)
CC DT 15-OCT-2001 (Rel. 40, Last annotation update)
CC DE Adenosyl-methionine-8-amino-7-oxonanoate aminotransferase
CC (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
CC aminotransferase) (Fragment).
CC BIOA.
CC GN Citrobacter freundii.
CC CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Citrobacter.
CC NCBI_TaxID=546;
CC !-
CC SEQUENCE FROM N.A.
CC MEDLINE=8906280; PubMed=2971595;
CC Shiuan D., Campbell A.;
CC "Transcriptional regulation and gene arrangement of Escherichia coli,
CC Citrobacter freundii and Salmonella typhimurium biotin operons.";
CC Gene 67:203-211(1988).
CC !- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-
CC oxonanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-
CC diaminonanoate.
CC !- COFACTOR: Pyridoxal phosphate.
CC !- PATHWAY: Biotin biosynthesis.
CC !- SUBUNIT: Homodimer.
```

```

CC -!- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21922; ; NOT ANNOTATED_CDS.
CC PIR; I40697; I40697.
CC InterPro; IPR0005814; Amirotrans 3.
CC PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.
CC Biotin biosynthesis; Transferase; Aminotransferase;
CC Pyridoxal phosphate.
CC NON TER 5
CC SEQUENCE 5 AA; 582 MW; 6AAAB1BA6F000C0 CRC64;
CC
CC Query Match 26.7%; Score 12; DB 1; Length 5;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+05;
CC Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 4 DD 5
CC ||
CC Db 4 DD 5
CC ||
CC
CC RESULT 6
CC TMOF SARBUB
CC ID TMOF SARBUB STANDARD; PRT; 6 AA.
CC AC P41435;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC ET 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 01-FEB-1996 (Rel. 33, Last annotation update)
CC DE Trypsin-modulating oostatic factor (TMOF).
CC OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
CC CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
CC CC Sarcophagidae; Sarcophaga.
CC NCBI_TaxID=7385;
CC [1]
CC RP SEQUENCE, AND SYNTHESIS.
CC RC TISSUE-Ovary;
CC RA MEDLINE=94211930; PubMed=8159807;
CC RA Bylemans D., Borovsky E., Hunt D.F., Shabanowitz J., Grauwels L.,
CC RA de Loof A.;
CC RT "Sequencing and characterization of trypsin modulating oostatic
CC RT factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria
CC RT (Sarcophaga) bullata.";
CC RL Regul. Pept. 50:61-72(1994).
CC CC -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS
CC CC IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN
CC CC CONCENTRATION IN THE HEMOLYPH RESULTING IN INHIBITION OF OOCYTE
CC CC DEVELOPMENT.
CC CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR
CC CC HORMONE.
CC KW Hormone.
CC SQ SEQUENCE 6 AA; 695 MW; 61E72451B76420C0 CRC64;
CC
CC Query Match 26.7%; Score 12; DB 1; Length 6;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+05;
CC Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 2 LH 3
CC ||
CC Db 5 LH 6
CC ||
CC
CC RESULT 7
CC UF03_MOUSE
CC ID UF03_MOUSE STANDARD; PRT; 7 AA.

```

```

AC P38641;
CC DT 01-OCT-1994 (Rel. 30, Created)
CC DT 01-OCT-1994 (Rel. 30, Last sequence update)
CC DE Unknown protein from 2D-page of fibroblasts (P36) (Fragment).
CC Mus musculus (Mouse).
CC CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
CC [1]
CC RN SEQUENCE.
CC RC T-SSUE=fibroblast;
CC RX MEDLINE=95009907; PubMed=7523108;
CC RA Verrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
CC RT "Separation and sequencing of familial and novel murine proteins
CC RT using preparative two-dimensional gel electrophoresis.";
CC RL Electrophoresis 15:735-745(1994).
CC CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC CC PROTEIN IS: 5.1, ITS MW IS: 36 kDa.
CC NON TER 7
CC SEQUENCE 7 AA; 842 MW; 6AA72B1DDB151180 CRC64;
CC
CC Query Match 26.7%; Score 12; DB 1; Length 7;
CC Best Local Similarity 33.3%; Pred. No. 1.3e+05;
CC Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 3 HDD 5
CC ||
CC Db - HEE 3
CC ||
CC
CC RESULT 8
CC NSKL SARBUB
CC ID NSKL SARBUB STANDARD; PRT; 9 AA.
CC AC P41492;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC ET 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 01-FEB-1996 (Rel. 33, Last annotation update)
CC DE Neosulfakinin-I (NEB-SK-I).
CC OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
CC CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
CC CC Sarcophagidae; Sarcophaga.
CC NCBI_TaxID=7385;
CC [1]
CC RP SEQUENCE.
CC RC T-SSUE=Head;
CC RX MEDLINE=93083101; PubMed=1360367;
CC RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.;
CC RT "Isolation and primary structure of two sulfakinin-like peptides from
CC RT the fleshfly, Neobellieria bullata.";
CC RL Comp. Biochem. Physiol. 103C:135-142(1992).
CC CC -!- FUNCTION: MYOTROPIC PEPTIDE.
CC CC -!- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
CC DR InterPro; IPR001651; Gastrin.
CC DR PROSITE; PS00259; Gastrin.
CC KW Neuropeptide; Amidation; Sulfation.
CC FT MOD_RES 4 4 SULFATION (POTENTIAL).
CC FT MOD_RES 9 9 AMIDATION (POTENTIAL).
CC SQ SEQUENCE 9 AA; 1187 MW; 8B0A0691E86B5AAA CRC64;
CC
CC Query Match 26.7%; Score 12; DB 1; Length 9;
CC Best Local Similarity 100.0%; Pred. No. 1.3e+05;
CC Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 4 DD 5
CC ||
CC Db 2 DD 3
CC ||
CC
CC RESULT 9
CC RE42_LITRU
CC ID RE42_LITRU STANDARD; PRT; 9 AA.

```





```
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRP/AMIDE RELATED PEPTIDE); 7
CC FAMILY
KW Neuropeptide; Amidation.
FT MOD RES 7
SQ SEQUENCE 7 AA; 992 MW; 69D407B5B11E350 CRC64;

Query Match 24.4%; Score 11; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 HDDL 6
Db 2 HEYL 5

RESULT 13
PTSP_BOMMO
ID PTSP_BOMMO STANDARD; PRT; 9 AA.
AC P820C3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Prothoracicostatic peptide (Bom-PTSP).
CS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
CX NCBI_TaxID=7091;
RN 1;
RP SEQUENCE.
RC STRAIN=C145 X N140; TISSUE=Brain;
RX MEDLINE=20002634; PubMed=10531108;
RA Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara M., Nagata S.,
RA Kataoka H.;
RT "Identification of a prothoracicostatic peptide in the larval brain of
RT the silkworm, Bombyx mori.";
RL J. Biol. Chem. 274:31169-31173 (1999).
RN 12;
RP ERRATUM.
RA Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara M., Nagata S.,
RA Kataoka H.;
RL J. Biol. Chem. 275:9892-9892 (2000).
CC -!- FUNCTION: Inhibits ecdysteroid biosynthesis in the prothoracic
CC gland.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: EARLY FIFTH INSTAR.
KW Hormone; Amidation.
FT MOD RES 9
SQ SEQUENCE 9 AA; 1090 MW; 3878C5B472AB6C3 CRC64;

Query Match 24.4%; Score 11; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 DELEA 9
Db 4 DLNSA 8

RESULT 14
TAL3_PICJA
ID TAL3_PICJA STANDARD; PRT; 9 AA.
AC P17441;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transaldolase III (EC 2.2.1.2) (Fragment).
OS Pichia jadinii (Yeast) (Candida utilis).
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Pichia.
CX NCBI_TaxID=4903;
RN 1;
RP SEQUENCE.
```

```
RX MEDLINE=75145197; PubMed=1092268;
RA Teclas O., Sun S.C.;
RT "Isolation of a peptide containing a histidyl-cysteiny sequence
RT from the active center of transaldolase.";
RL Arch. Biochem. Biophys. 167:525-533 (1975).
CC -!- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
CC -!- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde
CC 3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.
CC -!- PATHWAY: Pentose phosphate pathway; nonoxidative part.
CC -!- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY. SUBFAMILY 1.
DR PIR: A11497; A11497.
DR InterPro: IPR001585; Transaldolase.
DR PROSITE; PS00958; TRANSALDOLASE 2; PARTIAL.
DR PROSITE; PS01054; TRANSALDOLASE_1; PARTIAL.
KW Transferase; Pentose shunt.
FT NON_TER 1
FT NON_TER 9
SQ SEQUENCE 9 AA; 1033 MW; 325A31A44EB1E058 CRC64;

Query Match 24.4%; Score 11; DB 1; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.3e+05;
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LHDDL 7
Db 3 HCNLT 9

RESULT 15
ALL7_CVDPO
ID ALL7_CVDPO STANDARD; PRT; 7 AA.
AC P82156;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 7.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tetracoidea; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN 1;
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309 (1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD RES 7
SQ SEQUENCE 7 AA; 873 MW; 672879CABB569350 CRC64;

Query Match 22.2%; Score 10; DB 1; Length 7;
Best Local Similarity 33.3%; Pred. No. 1.3e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHD 4
Db 2 MYD 4

Search completed: October 20, 2003, 13:46:19
Job time : 13 secs
```

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 CompuGen Ltd.  
 OM protein - protein search, using sw model  
 Run on: October 20, 2003, 13:41:46 ; Search time 59 Seconds  
 (without alignments)  
 39.364 Million cell updates/sec

Title: SEQ1A  
 Perfect score: 45  
 Sequence: 1 vhlldllea 9

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 775

Minimum DB seq length: 0  
 Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SPTREMBL\_23:\*

1: sp\_archaea:  
 2: sp\_bacteria:  
 3: sp\_fungi:  
 4: sp\_human:  
 5: sp\_invertebrate:  
 6: sp\_mammal:  
 7: sp\_rnc:  
 8: sp\_organelle:  
 9: sp\_plant:  
 10: sp\_plant:  
 11: sp\_rodent:  
 12: sp\_virus:  
 13: sp\_vertebrate:  
 14: sp\_unclassified:  
 15: sp\_virus:  
 16: sp\_bacteriopl:  
 17: sp\_archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	19	42.2	9	2	Q47556
2	18	40.0	7	6	Q28742
3	16	35.6	7	8	Q91182
4	16	35.6	9	2	P82568
5	16	35.6	9	10	Q8LPT5
6	15	33.3	8	4	Q15900
7	15	33.3	9	9	Q9XUN0
8	14	31.1	8	2	P72279
9	14	31.1	8	2	Q93GR0
10	14	31.1	8	6	Q9BFA7
11	14	31.1	9	10	Q9PSZ2
12	14	31.1	9	11	C90Z28
13	14	31.1	9	11	Q8GC39
14	14	31.1	9	15	Q85710
15	14	31.1	9	15	Q8UTD7
16	13	28.9	7	12	Q66205

17 13 28.9 8 2 Q51594  
 18 13 28.9 8 3 Q9HDS4  
 19 13 28.9 8 4 Q15889  
 20 13 28.9 8 9 Q37854  
 21 13 28.9 9 2 Q8G126  
 22 12 26.7 7 10 Q49223  
 23 12 26.7 8 2 Q8KFX4  
 24 12 26.7 8 4 Q15894  
 25 12 26.7 8 4 Q16468  
 26 12 26.7 8 5 Q8MUN6  
 27 12 26.7 8 5 Q9NGM5  
 28 12 26.7 8 6 Q8WNS1  
 29 12 26.7 8 8 Q9GD00  
 30 12 26.7 8 11 Q9JLD7  
 31 12 26.7 8 13 Q9PS69  
 32 12 26.7 8 15 Q98YK9  
 33 12 26.7 9 2 Q43960  
 34 12 26.7 9 4 Q16605  
 35 12 26.7 9 4 Q9BQT4  
 36 12 26.7 9 11 Q61723  
 37 12 26.7 9 12 Q92766  
 38 12 26.7 9 13 Q8AYL5  
 39 12 26.7 9 13 Q8AUM7  
 40 11 24.4 7 2 Q8KMS9  
 41 11 24.4 7 2 Q07354  
 42 11 24.4 7 4 Q15897  
 43 11 24.4 8 3 Q05403  
 44 11 24.4 8 4 Q9JHK1  
 45 11 24.4 8 4 Q9P285

## ALIGNMENTS

## RESULT 1

Q47556  
 ID C47556 PRELIMINARY: PRT: 9 AA.  
 AC C47556  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created:  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE Aspartate transcarbamoylase regulatory chain (Fragment).  
 GN pyrB.  
 OS Escherichia coli.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID:562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82275957; PubMed=7051000;  
 RA Pazza C.D., Karels M.C., Navre M., Schachman H.K.;  
 RT "Genes encoding Escherichia coli aspartate transcarbamoylase: The  
 RT pyB-pyrB operon";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:4020-4024 (1982).  
 RN [2]  
 RP SEQUENCE OF 1-5 FROM N.A.  
 RX MEDLINE=83195079; PubMed=6302686;  
 RA Hoover T.A., Roof W.D., Foltermann K.F., O'Donovan G.A., Bencini D.A.,  
 RA Wild J.R.;  
 RT "Nucleotide sequence of the structural gene (pyrB) that encodes the  
 RT catalytic polypeptide of aspartate transcarbamoylase of Escherichia  
 RT coli";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2462-2466 (1983).  
 DR EMBJ; J01670; AAA24475.1; --  
 SQ SEQUENCE 9 AA; 1085 MW; 99EFD723344AA1F1 CRC64;  
 NON TER 9

Query Match 42.2%; Score 19; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 3 HDDLLE 8  
 ||| |:  
 ||| |:

3 HENKLE 8

SUITS 2  
8742

Q28742 PRELIMINARY: FRT; 7 AA.  
Q28743;  
C1-NOV-1996 (TEMBURE). 01, Created)  
C1-NOV-1996 (TEMBURE). 01, Last sequence update)  
C1-DEC-2001 (TEMBURE). 19, Last annotation update)  
Alpha-myosin heavy chain (Fragment).  
Cryotolagus cuniculus (Rabbit).  
Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
NCBI "taxid:9986;  
.1).  
SEQUENCE FROM N. A.  
MEDLINE=84221901; PubMed=6328491;  
Friedman D.J., Umeda P.K., Sinha A.M., Hsu H.-J., Jokovic S.,  
Rabinowitz M.;  
"Characterization of genomic clones specifying rabbit alpha- and beta-  
verteicular myosin heavy chains".  
Proc. Natl. Acad. Sci. U.S.A. 81:3044-3048(1984).  
EXBL; K01698; AAA31415.1; -.  
NON TER 1  
SEQUENCE 7 AA; 916 MW; 6B1B1AA1E69336B0 CRC64;

```
Query Match      40.0%; Score 18; DB 6; Length 7;
Best Local Similarity 50.0%; Pred. NO. 8.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Caps
```

2 LHDD 5  
: | :  
3 MHDE 6

```

RESULT: 3
099182 PRE-LIMINARY; PRT; 7 AA.
099182
01-MAY-1999 (TREMBLrel. 10, Created)
01-MAY-1999 (TREMBLrel. 10, Last sequence update)
01-DEC-2001 (TREMBLrel. 19, Last annotation update)
Cytochrome oxidase I (Fragment).
COI.
Gnatholebias zenatus.
Miteochondrion.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Cyprinodontiformes; Aplocheilidae; Rivulinae; Gnatholebias.
NCBI_taxid=135316;
[i]
SEQUENCE FROM N.A.
MEOLINE=20072928; PubMed=10603257;
Murphy W.J., Thomson J.E., Collier G.E.;
"Phylogeny of the Neotropical killifish family Rivulidae
(Cyprinodontiformes, Aplocheiloidei) inferred from mitochondrial DNA
sequences."
Mol. Phylogenet. Evol. 13:289-301(1999).
EMBL; AF002591; AAD01074.1; -.
Miteochondrion.
NON_TER 1
SEQUENCE 7 AA; 899 MW; 6727212FCB572030 CRC64;

```

Query Match 35.6%; Score 16; DB 8; Length 7;  
Best Local Similarity 42.9%; Pred. No. 8.3e+05;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps

```
1 VHSCDL 7  
   : |  
   ::|  
1 LYCHLL 7
```

```

      4
      4538
      PRELIMINARY;          9 AA.
      P82568
      01-OCT-2000 (TREM5Jrel. 15, Created)
      01-OCT-2000 (TREM5Jrel. 15, Last sequence update)
      01-OCT-2000 (TREM5Jrel. 15, Last annotation update)
      Unknown protein from 2D-page (Fragment).
      Streptococcus pyogenes.
      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
      Streptococcus.
      NCBI_TaxID=1314; .
      [.]
      SEQUENCE AND MASS SPECTROMETRY..
      STRAN=JES4;
      Hogan D.A., Du P., Stevenson T.I., Whitton M., Kirby G.W., Rogers J.,
      vanSogelen R.A.;
      "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
      proteins.";
      Submitted (MAY-2000) to the SWISS-PROT data bank.
      -- MASS SPECTROMETRY: YW=22592.04; METHOD=ELECTROSPRAY.
      NCN_TER      1
      NCN_TER      9
      SEQUENCE     9 AA; 1069 MW; 2A771042CB1AB2D7 CRC64;

```

```
Query Match      35.6%; Score 16; DB 2; Length 9;
Best Local Similarity 40.0%; Pred. NC. 8.3e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
4 DCLLE 8
      ...|
      ...|
      ...|
4 DEVIE 8
```

```

SUBJ 5
=====
LPT5
QBLP5 PRELIMINARY; PRT; 9 AA.
QBLP5:
01-OCT-2002 (REMBLrel. 22, Created)
01-OCT-2002 (REMBLrel. 22, Last sequence update)
01-OCT-2002 (REMBLrel. 22, Last annotation update)
Beta-expansin-like protein (Fragment).
Zea mays (Maize).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
PACCAD clade; Panicoideae; Andropogoneae; Zea.
NCBI_taxID=4577;
[.]
SEQUENCE FROM N.A.
STRAN=cy. C123;
Ching A.S., Caidwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,
Morgante M., Rafalski J.A.;
"SNP frequency, haplotype structure and linkage disequilibrium in
elite maize inbred lines";
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
EMBL: AF094310; AA021836.1; -.
NON_TER 1
SEQUENCE 9 AA; 977 NW; SC05B2D2CB1AAA3 CRC64;

```

```
Query Match: 35.6%; Score 16; DB 10; Length 9;
Best Local Similarity 33.3%; Pred.No. 8.3e+05;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```

4 DOLLEA 9
  :::|
4 DEVICA 9

```

SU-6  
5300  
Q15900  
PRELIMINARY;  
Q15900;  
PRT; 8 AA.

```

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (Clone XP7B1A) (Fragment)
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chirault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RC "Isolation of chromosome-specific genes by reciprocal probing of
R: arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L3279; AAA73890.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 931 MW; B5DDC403369AAEB; CRC64;

Query Match 33.3%; Score 15; DB 4; Length 8;
Best Local Similarity 42.9%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

CY 3 HDLLER 9
DB 1 HCDMKRA 7

RESULT 7
Q9XJNC PRELIMINARY; PRT; 9 AA.
ID Q9XJNC
AC Q9XJNC;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PIC (Fragment)
OS Bacteriophage phi-10.
OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
CX NCBI_TaxID=98889;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93350412; PubMed=0419946;
RA Mandich L., Qiao X., Qiao C., Onodera S., Romantschuk M.,
RA Hoogstraaten D.;
RC "Isolation of additional bacteriophages with genomes of segmented
R: double-stranded RNA.";
RL J. Bacteriol. 181:4505-4508(1999).
DR EMBL; AF125675; AA022555.1; -.
FT NON_TER 9 9
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1058 MW; 89CE376AA720544A; CRC64;

Query Match 33.3%; Score 15; DB 9; Length 9;
Best Local Similarity 40.0%; Pred. No. 8.3e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CY 4 DDILE 2
DB 2 DNILD 6

RESULT 8
P72279 PRELIMINARY; PRT; 8 AA.
ID P72279
AC P72279;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Biphenyl dioxygenase (Fragment)
GN BPHB.
OS Rhodococcus globerulus.

```

```

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
CX NCBI_TaxID=33008;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95253652; PubMed=7737502;
RA Asturias J.A., Diaz E., Timmis K.N.;
RT "Evolutionary relationship of the biphenyl dioxygenase of the gram-
RT positive bacterium Rhodococcus globerulus P6 to multicomponent
RT dioxygenases of gram-negative bacteria.";
RL Gene 156:11-18(1995).
DR EMBL; X80041; CAA56350.1; -.
KW Dioxygenase.
FT NON_TER 8 8
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 989 MW; EBD2C31AB6D73406; CRC64;

Query Match 31.1%; Score 14; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 8.3e+05;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

CY 2 LHDDLL 7
DB 3 LQDEVV 8

RESULT 9
Q93SRC PRELIMINARY; PRT; 8 AA.
ID Q93SRC
AC Q93SRC;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Beta-lactamase repressor BlaI (Fragment)
GN BLAI.
OS Staphylococcus epidermidis.
OG Plasmid p876.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
CX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6;
RA Sidhu M.S., Hair E., Sorum H., Holck A.L.;
RT "Genetic linkage between quaternary ammonium compound and beta-lactam
R: resistance in Staphylococci isolated from food.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF28779; AA038453.1; -.
KW Plasmid.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 930 MW; 4E3325B05AA44720; CRC64;

Query Match 31.1%; Score 14; DB 2; Length 8;
Best Local Similarity 50.3%; Pred. No. 8.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

CY 1 VLHD 4
DB 1 VLND 4

RESULT 10
Q9BFA7 PRELIMINARY; PRT; 8 AA.
ID Q9BFA7
AC Q9BFA7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment)
GN CREM.
OS Macroscelides proboscideus (Short-eared elephant shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Macroscelidea; Macroscelidae; Macroscelides.
CX NCBI_TaxID=29082;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=2108292; PubMed=11214319;
RA Murphy W., Ezirik E., Conson W.E., Zhang Y.P., Ryder O.A.,
O'Brien S.C.,
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL: AY01636; AAC47551.1; -.
FT NON TER 1 1
SQ SEQUENCE 8 AA; 940 MW; 2F1DD33AB5AB572A CRC64;

Query Match: 31.1%; Score 14; DB 6; Length 9;
Best Local Similarity 42.9%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LHDDLE 8
DB 2 LYSKAE 8

RESULT 11
Q9FSZ2 PRELIMINARY; PRT; 9 AA.
AC Q9FSZ2
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Hypothetical 1.0 kDa protein (Fragment).
CS Cicer arietinum (Chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OC NCBI_TaxID=3827;
RN 1;
RX STRAIN=cv. Castellana; TISSUE=Etiolated epicotyl;
RA Dopico B., Jimenez T., Labrador E.;
RT "cDNA clones expressed in etiolated Cicer arietinum epicotyls.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ299069; CAC10216.1; -.
AW Hypothetical protein.
FT NON TER 1 1
SQ SEQUENCE 9 AA; 990 MW; 9441BDDAA7272BEE CRC64;

Query Match: 31.1%; Score 14; DB 10; Length 9;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLEA 9
DB 4 LLDA 7

RESULT 12
Q9QZAB PRELIMINARY; PRT; 9 AA.
AC Q9QZAB
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE C-type lectin DCL1 (Fragment).
GN DCL1
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN 1;
RX GORSKI K., Huang X., Tseng S.-Y., Rattis F., Pardoll D., Tsuchiya H.;
RA "Dendritic cell regulation of DCL1 mRNA expression.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF192526; AAF04843.1; -.
AW MGD; MGI:2136650; Dcl1.
FT NON TER 1 1
SQ SEQUENCE FROM N.A.
RA GORSKI K., Huang X., Tseng S.-Y., Rattis F., Pardoll D., Tsuchiya H.;
RT "Dendritic cell regulation of DCL1 mRNA expression.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF192526; AAF04843.1; -.
AW MGD; MGI:2136650; Dcl1.
FT NON TER 1 1
SQ SEQUENCE 9 AA; 949 MW; 94AA144DDDD731AA CRC64;

Query Match: 31.1%; Score 14; DB 15; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DDLEA 9

```

Db : : :  
1 EDTLAA 6

RESULT 15

Q8UTD7  
ID Q8UTD7 PRELIMINARY; PRT; 9 AA.  
AC Q8UTD7;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Vpu protein.  
GN VPU.  
GN Human immunodeficiency virus 1.  
CC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
CX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=08W1471.27;  
RA Novitsky V.A., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,  
RA Williamson C., Ndong'u T., Klein I., Chang S.-Y., Peter T., Thior I.,  
RA Foley B.T., Gaoekwe S., Rybak N., Gasettswe S., Vamberg F.,  
RA Marlink R., Lee I.-H., Essex M.;  
RT "HIV-1 subtype C molecular phylogeny: consensus sequence for an AIDS  
RT vaccine design."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF443091; AAL34712.1;  
SQ SEQUENCE 9 AA; 1102 MW; 1888D40917272440 CRC64;

Query Match 31.1%; Score 14; DB 15; Length 9;  
Best Local Similarity 75.0%; Pred. NO. 8.3e-05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLE 9  
DB : : :  
3 NLE 6

Search completed: October 20, 2003, 13:48:25  
Job time : 60 secs

GenCore version: 5.1.6  
Copyright: (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:31:21 : Search time 56 Seconds  
(without alignments)  
25.510 Million cell updates/sec

Title: SEQ1A

Perfect score: 45

Sequence: 1 vlnhdlllea 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 179625

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:\*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb./AA196C.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1961.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb./AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb./AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb./AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb./AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb./AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARY-ES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	9	20	AAW97375
2	45	100.0	9	20	AAW99196
3	37	82.2	9	20	AAW97174
4	37	82.2	9	20	AAW99197
5	36	80.0	9	20	AAW97572
6	36	80.0	9	20	AAW99195
7	30	66.7	9	20	AAW97373
8	28	62.2	9	20	AAV10122
9	28	62.2	9	23	ABG79805

10	27	60.0	9	24	ABJ20115	MHC binding peptid
11	26	57.8	9	23	AAE31275	Human mages peptid
12	25	55.6	9	22	AAG79544	Amino terminal of
13	25	55.6	9	22	AA384492	Peptide fragment c
14	25	55.6	9	23	AAE26631	Yeast GPAl aminc t
15	24	53.3	6	13	AAE29358	Endothelin antagon
16	24	53.3	6	15	AAE69140	Endothelin C-termi
17	24	53.3	7	19	AAW33386	Altered -gm amino
18	24	53.3	7	22	AAW370589	Human immunoglobul
19	24	53.3	7	23	AAE28095	Human immunoglobul
20	24	53.3	7	23	AAO21066	Isomerised/optical
21	24	53.3	9	14	AAE43734	MHC Class I allele
22	24	53.3	9	14	AAE43735	MHC Class I allele
23	24	53.3	9	18	AAW39656	HPV16/18 E6 peptid
24	24	53.3	9	18	AAW39657	HPV16/18 E6 peptid
25	24	53.3	9	19	AAW54759	Peptide from HPV 1
26	24	53.3	9	20	AAV10561	HLA Class I motif
27	24	53.3	9	23	ABG90244	MHC class I molecu
28	23	51.1	6	19	AAW47424	Prenyl diphosphate
29	23	51.1	8	23	ABG93554	Human P-glycoprote
30	23	51.1	9	15	AAV38071	Hepatitis B virus-
31	23	51.1	9	20	AAV45639	Immunogenic peptid
32	23	51.1	9	23	AAU71211	Human MHC class I
33	23	51.1	9	23	AAU71428	Human MHC molecu
34	23	51.1	9	24	ABJ20107	MHC binding peptid
35	22	48.9	5	23	AAE71240	Human IgG1 catabol
36	22	48.9	5	23	AAU86928	Immunoglobulin cat
37	22	48.9	7	13	AAE28344	Escherichia coli D
38	22	48.9	7	13	AAE28346	Thermus aquaticus
39	22	48.9	7	21	AAE02941	Nucleotide-binding
40	22	48.9	7	22	AAE70588	Human immunoglobul
41	22	48.9	7	22	AAE70590	Pol 1 type DNA pol
42	22	48.9	7	23	ABG32949	Pol 1 type DNA pol
43	22	48.9	7	23	ABG32951	Isomerised/optical
44	22	48.9	7	23	AAO21065	Isomerised/optical
45	22	48.9	7	23	AAO21067	Isomerised/optical

#### ALIGNMENTS

RESULT :  
AAW97375  
ID AAW97375 standard; Protein; 9 AA.

XX AC AAW97375;

XX DT 13-MAY-1999 (first entry)

XX DE HA-1 H-allele sequence.

XX KW Intron; minor histocompatibility antigen HA-1; typing allele;  
XX H allele; R allele; polymorphic nucleotide; HA-1 typing;  
XX KW bone marrow transplant; severe aplastic anaemia; leukaemia;  
XX KW immune deficiency disease; ss.

XX CS Homo sapiens.

XX PN WC9905313-A2.

XX PD 04-FEB-1999.

XX PF 23-JUL-1999; 98WO-EP04928.

XX PR 02-JUN-1998; 98EP-0870125.

XX PR 23-JUL-1997; 97EP-0202303.

XX PA (JYU-E-) RIJKSUNIV LEIDEN.

XX PI Gou'lym E;

XX DR WPI; 1999-142960/12.

RESJLT 4  
AAW99197



```

ID AAW99197 standard; peptide; 9 AA.
XX
AC AAW99197;
XX
DT 20-MAY-1999 (first entry);
XX
DE Minor histocompatibility antigen HA-1 T-cell epitope #3.
XX
KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
KW diagnosis; aplastic anaemia; immune deficiency disease.
XX
CS Homo sapiens.
XX
PN WO9905174-A1.
XX
PD 04-FEB-1999.
XX
PF 23-JUL-1998; 98WO-NL00425.
XX
PR 23-JUL-1997; 97EP-0202303.
XX
PA (UYLE-) RIJKSUNIV LEIDEN.
XX
PI Engelhard VH, Goulmy EAJM, Hunt DF;
XX
DR WPI; 1999-153312/13.
XX
PT A new minor histocompatibility antigen, HA-1 - useful to treat
PT immune diseases and prevent rejection and host versus graft disease
PT in bone marrow and organ transplantation
XX
PS Disclosure; Page 15; 47pp; English.
XX
CC The present sequence represents a new peptide (P1) constituting a T-cell
CC epitope obtainable from the minor histocompatibility antigen HA-1. The
CC peptide is immunogenic and can be used as part of a vaccine. P1 is used
CC as a medicine, to induce tolerance for transplants, prevent rejection
CC and/or graft versus host disease, or to treat (auto) immune diseases.
CC In particular it can be used with bone marrow transplantation, in the
CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency
CC diseases.
XX
SQ Sequence 9 AA;
XX
Query Match 82.2%; Score 37; DB 20; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLKDDLEA 9
||| |||||
Db 1 VLKDDLEA 9

RESULT 5
AAW97572
ID AAW97572 standard; peptide; 9 AA.
XX
AC AAW97572;
XX
DT 20-MAY-1999 (first entry);
XX
DE T-cell epitope from the minor histocompatibility antigen HA-1.
XX
KW T-cell epitope; minor histocompatibility antigen HA-1; vaccine;
KW transplant rejection; Graft-versus-Host Disease; autoimmune disease;
KW neoplastic haematopoietic cell.
XX
CS Homo sapiens.
XX
PN
XX
Key Location/Qualifiers
FT Misc-difference 3 /note= "His or Arg"

```

```

XX WC9905173-A1.
XX
PD 04-FEB-1999.
XX
PF 23-JUL-1998; 98WO-NL00424.
XX
PR 23-JUL-1997; 97EP-0202303.
XX
PA (UYLE-) RIJKSUNIV LEIDEN.
XX
PI Engelhard VH, Goulmy EAJM, Hunt DF;
XX
DR WPI; 1999-142855/12.
XX
PT Immunogenic peptide from minor histocompatibility antigen HA-1 -
PT useful for inducing tolerance to transplants and prevent rejection
PT or graft-versus-host disease
XX
PS Claim 1; Page 39; 57pp; English.
XX
CC The present sequence represents an immunogenic peptide constituting a
CC T-cell epitope, obtainable from the minor histocompatibility antigen
CC HA-1. The peptide can be used in vaccines or pharmaceutical formulations
CC as medicines to induce tolerance for transplants so as to prevent
CC rejection and/or Graft-versus-Host Disease, or to treat autoimmune
CC diseases. Neoplastic haematopoietic cells presenting the peptides, in
CC an HLA class I context, can be eliminated after specific recognition
CC of the peptides. The peptides can also be used to raise antibodies,
CC T-cell receptor, B- and T-cells.
XX
SQ Sequence 9 AA;
XX
Query Match 80.0%; Score 36; DB 20; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLKDDLEA 9
||| |||||
Db 1 VLKDDLEA 9

RESULT 6
AAW99195
ID AAW99195 standard; peptide; 9 AA.
XX
AC AAW99195;
XX
DT 20-MAY-1999 (first entry);
XX
DE Minor histocompatibility antigen HA-1 T-cell epitope #1.
XX
KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
KW diagnosis; aplastic anaemia; immune deficiency disease.
XX
CS Homo sapiens.
XX
OS Synthetic.
XX
Key Location/Qualifiers
FT Misc-difference 3 /label= His, Arg
XX
PN WC9905174-A1.
XX
PD 04-FEB-1999.
XX
PR 23-JUL-1998; 98WO-NL00425.
XX
PR 23-JUL-1997; 97EP-0202303.
XX
PA (UYLE-) RIJKSUNIV LEIDEN.
XX

```

PI Engelhard VH, Goulmy EAJM, Hunt DF;  
 XX MPI: 1999-153312/13.  
 XX  
 PT A new minor histocompatibility antigen, HA-1 - useful to treat  
 PT immune diseases and prevent rejection and host versus graft disease  
 PT in bone marrow and organ transplantation  
 XX  
 XX Claim 1: Page 32; 47pp; English.  
 XX  
 CC The present sequence represents a new peptide (PI) constituting a T-cell  
 CC epitope obtainable from the minor histocompatibility antigen HA-1. The  
 CC peptide is immunogenic and can be used as part of a vaccine. PI is used  
 CC as a medicine, to induce tolerance for transplants, prevent rejection  
 CC and/or graft versus host disease, or to treat (auto) immune diseases.  
 CC In particular it can be used with bone marrow transplantation, in the  
 CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency  
 CC diseases.  
 XX  
 XX Sequence 9 AA;  
 XX  
 Query Match 80.0%; Score 36; DB 20; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VLHDDLLLEA 9  
 DB 1 VLXDDLLLEA 9  
 RESULT 7  
 AA97373  
 ID AA97373 standard; Peptide: 9 AA.  
 XX  
 AC AA97373;  
 XX  
 XX 13-MAY-1999 (first entry)  
 XX  
 DE Peptide epitope of HA-1 antigen.  
 XX  
 KW Intron; minor histocompatibility antigen HA-1; typing allele;  
 KW H allele; R allele; polymorphic nucleotide, HA-1 typing;  
 KW bone marrow transplant; severe aplastic anaemia; leukaemia;  
 KW immune deficiency disease.  
 XX  
 CS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT Misc-difference 2 /note= "not specified"  
 FT Misc-difference 6 /note= "not specified"  
 FT Misc-difference 7 /note= "not specified"  
 FT Misc-difference 7 /note= "not specified"  
 XX  
 EN W09905313-A2.  
 XX  
 PD 04-FEB-1999.  
 XX  
 XX 23-JUL-1998; 98WO-EP04928.  
 PF  
 XX 02-JUN-1998; 98EP-0870125.  
 PR 23-JUL-1997; 97EP-0202303.  
 XX  
 XX (UYLE-); RIJKSUNIV LEIDEN.  
 PA  
 PI Goulmy E;  
 XX  
 XX MPI; 1999-142960/12.  
 DR  
 XX Typing minor histocompatibility antigen HA-1 - by amplifying and  
 PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g.  
 PT detection of genetic aberrances

XX  
 PS Example 1: Page 22; 59pp; English.  
 XX  
 CC The present sequence represents an epitope of the minor  
 CC histocompatibility antigen HA-1. The specification describes  
 CC methods for typing alleles (preferably the H and R alleles) of  
 CC the minor histocompatibility antigen HA-1 in a sample, which comprise  
 CC detecting polymorphic nucleotides in the cDNA or genomic nucleic acids  
 CC of the alleles. The methods can be used for HA-1 typing for bone marrow  
 CC transplants, severe aplastic anaemia, leukaemia and immune deficiency  
 CC diseases, as well as detection of genetic aberrances. The probes and  
 CC primers of the invention can be used to screen for the HA-1 alleles.  
 CC The HA-1 peptides can be used anti-idiotypic B cells and/or T cells  
 CC and antibodies.  
 XX  
 XX Sequence 9 AA;  
 XX  
 Query Match 66.7%; Score 30; DB 20; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 9.3e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 VLHDDLLLEA 9  
 DB 1 VLXDDLLLEA 9  
 RESULT 8  
 AA970122  
 ID AA970122 standard; Peptide: 9 AA.  
 XX  
 AC AA970122;  
 XX  
 XX 12-MAY-1999 (first entry)  
 XX  
 DE T cell epitope/MHC ligand SEQ ID NO:52.  
 XX  
 KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;  
 KW immunisation; tumour; infectious disease; immunotherapy; cancer;  
 KW malignant melanoma; viral disease; hepatitis; AIDS.  
 XX  
 CS Synthetic.  
 CS Epstein-Barr virus.  
 XX  
 XX W09902183-A2.  
 XX  
 PD 21-JAN-1999.  
 XX  
 XX 10-JUL-1998; 98WO-US14289.  
 PF  
 XX 10-DEC-1997; 97US-0388320.  
 PR 10-JUL-1997; 97CA-2209815.  
 XX  
 PA (CTL-); CTL IMMUNOTHERAPIES CORP.  
 XX  
 PI Klendig TM, Simard JTL;  
 XX  
 DR WPI; 1999-120514/10.  
 XX  
 XX Inducing a cytotoxic T lymphocyte response - by maintaining a level  
 PT of antigen in the lymphatic system of a mammal so as to provide a  
 PT sustained CTL response, used to treat, e.g. AIDS  
 XX  
 XX Disclosure; Page 25; 199pp; English.  
 XX  
 CC The present invention describes a method of inducing and/or sustaining  
 CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The  
 CC method comprises: (a) delivering an antigen to the mammal at a level to  
 CC induce an immunological CTL response in the mammal; and (b) maintaining  
 CC the level of the antigen in the mammal's lymphatic system to maintain  
 CC the immunologic CTL response. The method can be used for the delivery of  
 CC e.g. a differentiation antigen, a tumour-specific multilineage antigen,  
 CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor  
 CC gene antigen, or a viral antigen. They can be used for the treatment of

CC disease such as cancer, e.g. malignant melanoma or infectious disease.  
 CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery  
 CC to the lymphatic system provides for potent CTL stimulation that takes  
 CC place in the milieu of the lymphoid organ, and it sustains stimulation  
 CC that is necessary to keep CTL active, cytotoxic and recirculating  
 CC through the body. AAY:0071 to AAY10633 represent examples of peptide  
 CC antigens given in the present invention.

XX  
 SQ Sequence 9 AA;  
 Query Match 62.2%; Score 28; DB 20; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDL 6  
 |||:  
 Db 4 VLHEDL 9

RESULT 9  
 ABG79805  
 ID ABG79805 standard; Peptide: 9 AA.  
 AC ABG79805;  
 DT 15-NOV-2002 (first entry)  
 DE MHC class I molecule, viral epitope #53.  
 KW Major histocompatibility complex; MHC; MHC class I molecule; virus;  
 KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system;  
 KW antigen; immunogenic; malignant tumour; carcinoma; melanoma; leukaemia;  
 KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis;  
 KW acquired immune deficiency syndrome; AIDS.  
 OS Epstein Barr virus.  
 PN WO200262368-A2.  
 XX 15-AUG-2002.  
 PF 22-JAN-2002; 2002WO-US02033.  
 PR 02-FEB-2003; 2001US-0776232.  
 PA (CTLI-) CTL IMMUNOTHERAPIES CORP.  
 PI Kundig TM, Simard JOL;  
 PS WPI; 2002-657506/70.  
 XX  
 PT Inducing or sustaining immunological cytotoxic T lymphocyte response in  
 PT a mammal, useful for treating a mammal with malignant tumour or  
 PT infectious disease, by directly administering an antigen to the  
 PT lymphatic system of the mammal.  
 XX  
 PS Disclosure; Page 19; 73pp; English.  
 CC The invention relates to a method of inducing and/or sustaining an  
 CC immunological cytotoxic T lymphocyte (CTL) response in a mammal  
 CC comprising administering directly to the lymphatic system of the mammal:  
 CC (a) an antigen in the form of a polypeptide; (b) a vector comprising a  
 CC nucleic acid encoding the antigen; or (c) a non-peptide antigen. The  
 CC method is useful for inducing and/or sustaining CTL response in a mammal.  
 CC This is particularly useful for treating a mammal having a malignant  
 CC tumour (e.g. carcinoma, melanoma, leukaemia or lymphoma) or infectious  
 CC disease (e.g. hepatitis, acquired immune deficiency syndrome (AIDS),  
 CC malaria, measles or tuberculosis), or in an animal having a  
 CC predisposition to these diseases. The mammal may be dogs, cats, mice,  
 CC cattle, sheep, pigs, goats, rabbits, or preferably humans.  
 CC ABG79753-ABG80319 represent viral epitopes on major histocompatibility  
 CC complex (MHC) class I molecules, used in the method of the invention.

SQ Sequence 9 AA;  
 Query Match 62.2%; Score 28; DB 23; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDDL 6  
 |||:  
 Db 4 VLHEDL 9

RESULT 10  
 ABJ20115  
 ID ABJ20115 standard; Peptide: 9 AA.  
 AC ABJ20115;  
 DT 10-APR-2003 (first entry)  
 DE MHC binding peptide SEQ ID No 28C.  
 XX Antirheumatic; anti-allergic; antiarthritic; neurotropic; neuroprotective;  
 KW anti-inflammatory; major histocompatibility complex; MHC;  
 KW autoimmune disease; T cell; B cell; allergic disease; multiple sclerosis;  
 KW rheumatoid arthritis; neurodegenerative disorder; Alzheimer's disease;  
 KW inflammation; gene therapy; MHC binding peptide.  
 OS Synthetic.  
 PN WO200294981-A2.  
 XX 28-NOV-2002.  
 PF 16-MAY-2002; 2002WO-IL00383.  
 PR 16-MAY-2001; 2001US-290958P.  
 PR 29-MAY-2001; 2001US-0865548.  
 XX (TECR ) TECHNION RES & DEV FOUND LTD.  
 XX Barnea E, Beer I, Ziv T, Admon A, Dassau L, Buchsbaum S;  
 XX WPI; 2003-210043/20.  
 XX  
 PT Identifying peptides that are capable of binding to major  
 PT histocompatibility complex (MHC) molecules of a particular haplotype by  
 PT analyzing peptides bound to the soluble and secreted form of the MHC  
 PT molecules of the particular haplotype.  
 XX  
 PS Claim 59; Page 215; 238pp; English.  
 CC The invention relates to a novel method for identifying peptides  
 CC originating from a particular cell type, which are capable of binding to  
 CC major histocompatibility complex (MHC) molecules of a particular  
 CC haplotype. The method comprises analysing peptides bound to the soluble  
 CC and secreted form of the MHC molecules of the particular haplotype. The  
 CC method is useful for identifying peptides for treating an autoimmune  
 CC disease, such as T or B cell and/or allergic disease or condition,  
 CC rheumatoid arthritis, or multiple sclerosis, neurodegenerative disorders,  
 CC e.g. Alzheimer's disease, or diseases associated with inflammation. The  
 CC sequences of the invention may be used in a gene therapy application.  
 CC This sequence represents a peptide relating to the method for identifying  
 CC MHC binding peptides of the invention.

XX  
 SQ Sequence 9 AA;  
 Query Match 60.0%; Score 27; DB 24; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDD 5  
 |||:  
 Db 1 ILHDD 5





XX The peptide is an endothelin antagonist useful in controlling  
CC hypertension, myocardial infarction, congestive heart failure,  
CC endotoxic shock, subarachnoid haemorrhage, asthma, arrhythmias,  
CC acute renal failure, pre-eclampsia, diabetes and metabolic,  
CC endocrinological and neurological disorders. Administration is oral  
CC parenteral or by inhalation in doses of 0.01-20 (esp. 0.01-10) mg/kg/  
CC day. It may be prepared by conventional peptide synthesis.  
CC (Updated on 25-Mar-2003 to correct PN field.)

XX SQ Sequence 6 AA;

Query Match 53.3%; Score 24; DB 13; Length 6;  
Best Local Similarity 60.0%; Pred. No. 9.3e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 3;

OY 3 HDLL 7  
| | | | |  
Db 1 HDII 5

Search completed: October 20, 2003, 13:45:47  
Job time : 58 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OX protein - protein search, using sw model

Run on: October 20, 2003, 13:48:32 ; Search time 43 Seconds  
(without alignments)  
34.308 Million cell updates/sec

Title: SEQLA  
Perfect score: 45  
Sequence: 1 vlhddlea 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 609560 seqs, 163917102 residues

Total number of hits satisfying chosen parameters: 60600

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOVB.pcp:\*
- 2: /cgn2\_6/ptodata/2/pubaa/PCT\_NEW\_PUB.pcp:\*
- 3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pcp:\*
- 4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOVB.pcp:\*
- 5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pcp:\*
- 6: /cgn2\_6/ptodata/2/pubaa/PCTUS\_PUBCOVB.pcp:\*
- 7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pcp:\*
- 8: /cgn2\_6/ptodata/2/pubaa/US08\_PUBCOVB.pcp:\*
- 9: /cgn2\_6/ptodata/2/pubaa/US09A\_PUBCOVB.pcp:\*
- 10: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOVB.pcp:\*
- 11: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOVB.pcp:\*
- 12: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB.pcp:\*
- 13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOVB.pcp:\*
- 14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOVB.pcp:\*
- 15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOVB.pcp:\*
- 16: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB.pcp:\*
- 17: /cgn2\_6/ptodata/2/pubaa/US0C\_NEW\_PUB.pcp:\*
- 18: /cgn2\_6/ptodata/2/pubaa/US0C\_PUBCOVB.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	57.8	6	12 US-10-166-225A-163	Sequence 163, App
2	26	57.8	6	12 US-10-166-225A-164	Sequence 164, App
3	26	57.8	6	12 US-10-166-225A-165	Sequence 165, App
4	26	57.8	6	12 US-10-166-225A-166	Sequence 166, App
5	25	55.6	6	12 US-10-166-225A-167	Sequence 167, App
6	25	55.6	6	12 US-10-166-225A-168	Sequence 168, App
7	25	55.6	6	11 US-09-953-354-3	Sequence 3, Appli
8	25	55.6	9	15 US-10-247-074-14	Sequence 14, Appl
9	25	55.6	9	15 US-10-247-074-16	Sequence 16, Appl
10	24	53.3	7	12 US-10-020-354-86	Sequence 86, Appl
11	23	51.1	8	10 US-09-982-772-74	Sequence 74, Appl
12	23	51.1	9	9 US-09-834-765-132	Sequence 132, App
13	23	51.1	9	9 US-09-834-765-349	Sequence 349, App
14	22	48.9	5	9 US-09-933-497B-30	Sequence 30, Appl
15	22	48.9	6	10 US-09-911-838-156	Sequence 156, App

15	22	48.9	7	10 US-09-911-838-155	Sequence 155, App
17	22	48.9	7	10 US-09-911-838-157	Sequence 157, App
18	22	48.9	9	15 US-10-001-469-24	Sequence 24, Appl
19	22	48.9	9	15 US-10-001-469-26	Sequence 26, Appl
20	22	48.9	9	15 US-10-001-469-215	Sequence 215, App
21	22	48.9	9	15 US-10-001-469-416	Sequence 416, App
22	22	48.9	9	15 US-10-001-469-453	Sequence 453, App
23	22	48.9	9	15 US-10-001-469-608	Sequence 608, App
24	22	48.9	9	15 US-10-001-469-668	Sequence 668, App
25	22	48.9	9	15 US-10-001-469-842	Sequence 842, App
26	22	48.9	9	15 US-10-001-469-1222	Sequence 1222, Ap
27	22	48.9	9	15 US-10-001-469-1246	Sequence 1246, Ap
28	22	48.9	9	15 US-10-001-469-1470	Sequence 1470, Ap
29	22	48.9	9	15 US-10-001-469-1586	Sequence 1586, Ap
30	22	48.9	9	15 US-10-001-469-1609	Sequence 1609, Ap
31	22	48.9	9	15 US-10-001-469-1716	Sequence 1716, Ap
32	22	48.9	9	15 US-10-001-469-1805	Sequence 1805, Ap
33	22	48.9	9	15 US-10-001-469-1829	Sequence 1829, Ap
34	22	48.9	9	15 US-10-001-469-1876	Sequence 1876, Ap
35	22	48.9	9	15 US-10-001-469-1877	Sequence 1877, Ap
36	22	48.9	9	15 US-10-001-469-1920	Sequence 1920, Ap
37	22	48.9	9	15 US-10-001-469-1957	Sequence 1957, Ap
38	22	48.9	9	15 US-10-001-469-2052	Sequence 2052, Ap
39	21	46.7	9	10 US-09-918-243-76	Sequence 76, Appl
40	21	46.7	9	10 US-09-905-083-76	Sequence 76, Appl
41	21	46.7	9	12 US-09-833-203-54	Sequence 54, Appl
42	21	46.7	9	12 US-10-144-188-34	Sequence 34, Appl
43	21	46.7	9	12 US-10-200-708-656	Sequence 656, App
44	21	46.7	9	15 US-10-001-469-696	Sequence 696, App
45	21	46.7	9	15 US-10-001-469-895	Sequence 895, App

## ALIGNMENTS

RESULT 1  
US-10-166-225A-163  
; Sequence 163, Application US/10166225A  
; Publication No. US030148416A1  
; GENERAL INFORMATION:  
; APPLICANT: BERRY, Alan  
; APPLICANT: BRETZEL, Werner  
; APPLICANT: HUMBERLIN, Markus  
; APPLICANT: LOPEZ-ULIBARRI, Rual  
; APPLICANT: MAYER, Arne F.  
; APPLICANT: VELISSEV, Alexei A.  
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION  
; FILE REFERENCE: C38435/121966  
; CURRENT APPLICATION NUMBER: US/10166,225A  
; CURRENT FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 163  
; LENGTH: 6  
; TYPE: PKC  
; ORGANISM: Bacillus stearothermophilus  
US-10-166-225A-163

Query Match 57.8; Score 26; DB 12; Length 6;  
Best Local Similarity 80.0; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHDDJ 6  
DB 1 LHDDJ 5

RESULT 2  
US-10-166-225A-164  
; Sequence 164, Application US/10166225A  
; Publication No. US20030148416A1  
; GENERAL INFORMATION:  
; APPLICANT: BERRY, Alan

```

; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: YELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/10/166,225A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-166-225A-164

Query Match          57.8%; Score 26; DB 12; Length 6;
Best Local Similarity 80.0%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDDL 6
Db 1 IHDDL 5

RESULT 3
US-10-166-225A-165
; Sequence 165, Application US/10166225A
; Publication No. US20030148416A1
; GENERAL INFORMATION:
; APPLICANT: BERRY, Alan
; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: YELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/10/166,225A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 165
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-166-225A-165

```

```

Query Match          57.8%; Score 26; DB 12; Length 6;
Best Local Similarity 80.0%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDDL 6
Db 1 IHDDL 5

RESULT 4
US-10-166-225A-166
; Sequence 166, Application US/10166225A
; Publication No. US20030148416A1
; GENERAL INFORMATION:
; APPLICANT: BERRY, Alan
; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: YELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/10/166,225A
; CURRENT FILING DATE: 2002-06-05

```

```

; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 165
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-166-225A-166

Query Match          57.8%; Score 26; DB 12; Length 6;
Best Local Similarity 80.0%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDDL 6
Db 1 IHDDL 5

RESULT 5
US-10-166-225A-161
; Sequence 161, Application US/10166225A
; Publication No. US20030148416A1
; GENERAL INFORMATION:
; APPLICANT: BERRY, Alan
; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: YELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/10/166,225A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 161
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Bradyrhizobium japonicum
US-10-166-225A-161

Query Match          55.6%; Score 25; DB 12; Length 6;
Best Local Similarity 80.0%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDDL 6
Db 1 VHDDL 5

```

```

RESULT 6
US-10-166-225A-162
; Sequence 162, Application US/10166225A
; Publication No. US20030148416A1
; GENERAL INFORMATION:
; APPLICANT: BERRY, Alan
; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: YELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/10/166,225A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 162
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Rhizobium sp. strain NGR234
US-10-166-225A-162

Query Match          55.6%; Score 25; DB 12; Length 6;

```



```
Best Local Similarity 80.0%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHDDL 6
Db 1 VHDDL 5

RESULT 7
US-09-953-354-3
; Sequence 3, Application US/09953354
; Publication No. US20030054402A1
; GENERAL INFORMATION:
; APPLICANT: Klein, Christine A.
; APPLICANT: Murphy, Andrew J. M.
; TITLE OF INVENTION: Methods and Compositions for
; Identifying Receptor Effectors
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSES: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/953,354
; FILING DATE: 13-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/689,172
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: KARA, Catherine J.
; REGISTRATION NUMBER: 041,106
; REFERENCE/DOCKET NUMBER: CPI-012CP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-953-354-3

Query Match 55.6%; Score 25; DB 1; Length 9;
Best Local Similarity 33.3%; Pred. No. 5.4e+05;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 VHDDLLEA 9
Db 1 LIHEDIKA 9

RESULT 8
US-10-267-074-14
; Sequence 14, Application US/10267074
; Publication No. US20030108999A1
; GENERAL INFORMATION:
; APPLICANT: Manfredi, John
; APPLICANT: Benton, Benjamin K.
; APPLICANT: Wu, Meny-Yu
; TITLE OF INVENTION: CELLS HAVING AMPLIFIED SIGNAL TRANSDUCTION PATHWAY
; CURRENT APPLICATION NUMBER: US/10/020,354
; FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/254,884

Best Local Similarity 80.0%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHDDL 6
Db 1 VHDDL 5

RESULT 9
US-10-267-074-16
; Sequence 16, Application US/10267074
; Publication No. US20030108999A1
; GENERAL INFORMATION:
; APPLICANT: Manfredi, John
; APPLICANT: Benton, Benjamin K.
; APPLICANT: Wu, Meny-Yu
; TITLE OF INVENTION: CELLS HAVING AMPLIFIED SIGNAL TRANSDUCTION PATHWAY
; CURRENT APPLICATION NUMBER: US/10/267,074
; FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: US/09/378,046
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CONSENSUS MOTIF
US-10-267-074-16

Query Match 55.6%; Score 25; DB 15; Length 9;
Best Local Similarity 33.3%; Pred. No. 5.4e+05;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 VHDDLLEA 9
Db 1 LIHEDIKA 9

RESULT 10
US-10-020-354-96
; Sequence 86, Application US/10020354
; Publication No. US20030108999A1
; GENERAL INFORMATION:
; APPLICANT: DALL'ACQUA, WILLIAM
; APPLICANT: JOHNSON, LESLIE
; APPLICANT: WARD, ELIZABETH SALLY
; TITLE OF INVENTION: MOLECULES WITH EXTENDED HALF-LIVES, COMPOSITIONS AND USES THERE
; FILE REFERENCE: 10271-027
; CURRENT APPLICATION NUMBER: US/10/020,354
; FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/254,884
```

PRIOR FILING DATE: 2000-12-12  
PRIOR APPLICATION NUMBER: 60/238,760  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 118  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 86  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-020-354-86

Query Match 53.3%; Score 24; DB 12; Length 7;  
Best Local Similarity 71.4%; Pred. No. 5.4e+05;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLHDDL 7  
DB 1 VLHQDWL 7

RESULT 11  
US-09-982-172-74  
Sequence 74, Application US/C9982172  
Patent No. US2002013119A1

GENERAL INFORMATION:  
APPLICANT: Emil Israel Katz  
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
TITLE OF INVENTION: JMWZING EACH  
FILE REFERENCE: C1/22283  
CURRENT APPLICATION NUMBER: US/C9/982,172  
CURRENT FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 74  
LENGTH: 5

TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Computer generated synthetic peptide

US-09-982-172-74

Query Match 51.1%; Score 23; DB 10; Length 8;  
Best Local Similarity 50.0%; Pred. No. 5.4e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 HDLLE 8  
DB 3 HDELK 8

RESULT 12  
US-09-834-765-132  
Sequence 132, Application US/09834765  
Patent No. US20020055478A1

GENERAL INFORMATION:  
APPLICANT: Mary Paris  
APPLICANT: Pia M. Chailita-Eid  
APPLICANT: Arthur B. Raitano  
APPLICANT: Steve Chappell Mitchell  
APPLICANT: Daniel E.H. Afar  
APPLICANT: Aya Jakobovits  
TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
TITLE OF INVENTION: AND DETECTION OF CANCER  
FILE REFERENCE: 129.6USU1  
CURRENT APPLICATION NUMBER: US/C9/834,765  
CURRENT FILING DATE: 2001-09-21  
PRIOR APPLICATION NUMBER: 60/197,647  
PRIOR FILING DATE: 2000-04-12  
NUMBER OF SEQ ID NOS: 770  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 132  
LENGTH: 9

Query Match 48.9%; Score 22; DB 9; Length 5;  
Best Local Similarity 80.2%; Pred. No. 5.4e+05;

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-834-765-132

Query Match 51.1%; Score 23; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.4e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLEA 9  
DB 1 DLLEA 5

RESULT 13

US-09-834-765-349  
Sequence 349, Application US/09834765  
Patent No. US20020055478A1  
GENERAL INFORMATION:  
APPLICANT: Mary Paris  
APPLICANT: Pia M. Chailita-Eid  
APPLICANT: Arthur B. Raitano  
APPLICANT: Steve Chappell Mitchell  
APPLICANT: Daniel E.H. Afar  
APPLICANT: Aya Jakobovits

TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
TITLE OF INVENTION: AND DETECTION OF CANCER  
FILE REFERENCE: 129.6USU1  
CURRENT APPLICATION NUMBER: US/09/834,765  
CURRENT FILING DATE: 2001-09-21  
PRIOR APPLICATION NUMBER: 60/197,647  
PRIOR FILING DATE: 2000-04-12  
NUMBER OF SEQ ID NOS: 770  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 349  
LENGTH: 9

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-834-765-349

Query Match 51.1%; Score 23; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.4e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLEA 9  
DB 5 DLLEA 9

RESULT 14

US-09-933-497B-30  
Sequence 30, Application US/0933497B  
Patent No. US20020094192A1

GENERAL INFORMATION:  
APPLICANT: Ward, Elizabeth S.  
TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAINS WITH INCREASED HALF LIVES  
FILE REFERENCE: JUSC1483  
CURRENT APPLICATION NUMBER: US/09/933,497B  
CURRENT FILING DATE: 2002-02-07  
PRIOR APPLICATION NUMBER: 60/013,563  
PRIOR FILING DATE: 1996-03-18  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 30  
LENGTH: 5

TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptides

US-09-933-497B-30

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHDD 5  
| | | |  
Dc 1 VLHQD 5

## RESULT 15

US-09-911-838-156  
; Sequence 156, Application US/099-1838  
; Patent No. US20020151678A;  
; GENERAL INFORMATION:  
; APPLICANT: ARLINGHAUS, RALPH  
; TITLE OF INVENTION: PROPHYLAXIS AND THERAPY OF ACQUIRED IMMUNODEFICIENCY  
; FILE REFERENCE: US-09-911-838  
; CURRENT APPLICATION NUMBER: US/09/911-838  
; CURRENT FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 07/834,923  
; PRIOR FILING DATE: 1992-02-13  
; NUMBER OF SEQ ID NOS: 226  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 156  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-911-838-156

Query Match 48.8%; Score 22; DB 10; Length 6;  
Best Local Similarity 33.3%; Pred. No. 5.4e+05;  
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LHDJLL 7  
: | : | :  
Db 1 MHEDI 6

Search completed: October 20, 2003, 14:01:22  
Cob time : 44 secs



QY 1 VLHDD 5  
DB 3 LHHDD 7

## RESULT 2

US-09-808-126-5  
; Sequence 5, Application US/09808126  
; Patent No. 6410280  
; GENERAL INFORMATION:  
; APPLICANT: Obata, Shusei  
; Nishino, Tokuzo  
; Koyama, Tanetoshi  
; Sato, Yoshihiro  
; TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: KENYON & KENYON  
; STREET: 1500 K Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/808,126  
FILING DATE: 08-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/025,819  
ATTORNEY/AGENT INFORMATION:  
NAME: Khalilian, Hour  
REGISTRATION NUMBER: 39,546  
REFERENCE/DOCKET NUMBER: 10235/2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-220-4200  
TELEFAX: 202-220-4200  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-808-126-5  
Query Match 55.6%; Score 25; DB 4; Length 7;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDD 5  
DB 3 LHHDD 7

## RESULT 3

US-09-803-951-5  
; Sequence 5, Application US/09803951  
; Patent No. 6413761  
; GENERAL INFORMATION:  
; APPLICANT: Obata, Shusei  
; Nishino, Tokuzo  
; Koyama, Tanetoshi  
; Sato, Yoshihiro  
; TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: KENYON & KENYON  
STREET: 1500 K Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/803,951  
FILING DATE: 13-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/025,819  
ATTORNEY/AGENT INFORMATION:  
NAME: Khalilian, Hour  
REGISTRATION NUMBER: 39,546  
REFERENCE/DOCKET NUMBER: 10235/2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-220-4200  
TELEFAX: 202-220-4200  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-803-951-5  
Query Match 55.8%; Score 25; DB 4; Length 7;  
Best Local Similarity 80.3%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDD 5  
DB 3 LHHDD 7

## RESULT 4

US-08-582-333A-3  
; Sequence 3, Application US/08582333A  
; Patent No. 6255059  
; GENERAL INFORMATION:  
; APPLICANT: Klein, Christine A.  
; TITLE OF INVENTION: Methods and Compositions for  
; IDENTIFYING RECEPTOR EFFECTORS  
; NUMBER OF SEQUENCES: 98  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Ascii(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/582,333A  
FILING DATE: 17-JAN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Catherine J. Kara  
REGISTRATION NUMBER: 41,106

REFERENCE/DOCKET NUMBER: CPI-012CP5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-582-333A-3

Query Match 55.6%; Score 25; DB 3; Length 9;  
Best Local Similarity 33.3%; Pred. No. 2.5e+05;  
Matches 3; Conservative 5; Mismatches 1; Indels 0;

QY 1 VLHDDLEA 9  
: : : : :  
DB 1 LIHEDIAXA 9

RESULT 5  
US-09-305-923A-7  
Sequence 7, Application US/09305923A  
Patent No. 6355473  
GENERAL INFORMATION:  
APPLICANT: Ostanin, Kiril  
APPLICANT: Silverman, Lauren  
TITLE OF INVENTION: YEAST CELLS HAVING MUTATIONS IN *stp22* AND USES THEREFOR  
FILE REFERENCE: CPI-091  
CURRENT APPLICATION NUMBER: US/09/305,923A  
CURRENT FILING DATE: 1999-05-05  
PRIOR APPLICATION NUMBER: 60/084,420  
PRIOR FILING DATE: 1998-05-06  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 9  
TYPE: PRT  
ORGANISM: *Saccharomyces cerevisiae*  
US-09-305-923A-7

Query Match 55.6%; Score 25; DB 4; Length 9;  
Best Local Similarity 33.3%; Pred. No. 2.5e+05;  
Matches 3; Conservative 5; Mismatches 1; Indels 0;

QY 1 VLHDDLEA 9  
: : : : :  
DB 1 LIHEDIAXA 9

RESULT 6  
US-09-217-609A-8  
Sequence 8, Application US/09217609A  
Patent No. 6071733  
GENERAL INFORMATION:  
APPLICANT: MURAMATSU, Masayoshi  
APPLICANT: KOIKE, Ayumi  
APPLICANT: OGURA, Kyoza  
APPLICANT: KOYAMA, Tabetoshi  
APPLICANT: SHIMIZU, Naoto  
APPLICANT: CHO, Yenwin  
TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESS: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, NW - Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: US  
ZIP: 20036  
COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: WordPerfect 6.1 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/217,609A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/873,235  
FILING DATE: 11-Jun-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: TOFFENETTI, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 10235/1  
TELEPHONE: 202-429-1776  
TELEFAX: 202-429-0736  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-217-609A-8  
Query Match 51.1%; Score 23; DB 3; Length 6;  
Best Local Similarity 60.0%; Pred. No. 2.5e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0;

QY 1 VLHDD 5  
: : : : :  
DB 2 LIHDD 6

RESULT 7  
US-08-873-235B-8  
Sequence 8, Application US/08873235B  
Patent No. 6174715  
GENERAL INFORMATION:  
APPLICANT: MURAMATSU, Masayoshi  
APPLICANT: KOIKE, Ayumi  
APPLICANT: OGURA, Kyoza  
APPLICANT: KOYAMA, Tabetoshi  
APPLICANT: SHIMIZU, Naoto  
APPLICANT: CHO, Yenwin  
TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESS: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, NW - Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: US  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: WordPerfect 6.1 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/873,235B  
FILING DATE: 11-Jun-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 154441/1996  
FILING DATE: 14-Jun-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: TOFFENETTI, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 10235/1  
TELEPHONE: 202-429-1776

TELEFAX: 202-429-0796  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-873-2359-8

Query Match 51.1%; Score 23; DB 3; Length 6;  
Best Local Similarity 60.0%; Pred. No. 2.5e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHDD 5  
DB 2 LIHDD 6

RESULT 8  
US-08-159-339A-370  
Sequence 370, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:

APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esteban  
TITLE OF INVENTION: HLA Binding peptides and Their  
TITLE OF INVENTION: Uses  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA

ZIP: 94111-3934  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/159.339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424

Prior APPLICATION DATA:

APPLICATION NUMBER: US 07/326,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993

APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Ellen Laufer  
REGISTRATION NUMBER: 32,762

REFERENCE/DOCKET NUMBER: 018623-005030US  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

TELEX:

INFORMATION FOR SEQ ID NO: 370:

SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-159-339A-370

Query Match 51.1%; Score 23; DB 3; Length 9;  
Best Local Similarity 42.9%; Pred. No. 2.5e+05;  
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LHDDLE 8  
DB 2 LHDDLE 8

RESULT 9

US-08-811-463-30  
Sequence 30, Application US/08811463C  
Patent No. 6277375  
GENERAL INFORMATION:

APPLICANT: Ward, Elizabeth S.  
TITLE OF INVENTION: IMMUNOGLOBULIN-LIKE DOMAINS WITH INCREASED HALF LIVES  
FILE REFERENCE: UTSD:483  
CURRENT APPLICATION NUMBER: US/08/811.463C  
CURRENT FILING DATE: 1997-03-03  
EARLIER APPLICATION NUMBER: 60/013,563  
EARLIER FILING DATE: 1996-03-18  
NUMBER OF SEQ ID NOS: 39

SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 30

LENGTH: 5  
TYPE: PRT

ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptides  
US-08-811-463-30

Query Match 48.9%; Score 22; DB 3; Length 5;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLHDD 5  
DB 1 VLHDD 5

RESULT 10

US-08-208-036-7

Sequence 7, Application US/08208036  
Patent No. 5436326  
GENERAL INFORMATION:

APPLICANT: Yoshizumi, ICHINO et al.  
TITLE OF INVENTION: METHOD FOR CLONING OF A GENE FOR PCL I TYPE

TITLE OF INVENTION: DNA POLYMERASE  
NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wengert, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington  
STATE: D.C.

COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DCS  
SOFTWARE: DisplayWrite

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/208,036  
FILING DATE:

CLASSIFICATION: 435  
Prior APPLICATION DATA:

APPLICATION NUMBER: US/07/887,282  
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850  
TELEFAX:

```

;
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 7 amino acid residues
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLES:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
;   NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
;
; US-08-208-036-7

```

```

Query Match: 48.9%; Score 22; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 LHDDL 7
DB 2 VHDELV 7

```

```

RESULT 11
US-08-208-036-9
; Sequence 9, Application JS/08208036
; Patent No. 5436326
; GENERAL INFORMATION:
; APPLICANT: Yoshizumi ISHINO et al.
; TITLE OF INVENTION: METHOD FOR CLONING OF A GENE FOR POL I TYPE
; TITLE OF INVENTION: DNA POLYMERASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:

```

```

; MEDICAL TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Displaywrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,036
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/887,282
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,357
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 7 amino acid residues
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: Thermus aquaticus
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLES:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
;   NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
;
; US-08-208-036-9

```

```

Query Match: 49.9%; Score 22; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHDDL 7
DB 2 VHDELV 7

```



RESULT 12  
US-08-428-823-7  
Sequence 7, Application US/08428823  
Patent No. 5753482  
GENERAL INFORMATION:  
APPLICANT: Yoshizumi ISHINO et al.  
TITLE OF INVENTION: METHOD FOR CLONING OF A GENE FOR  
POLYMERASE  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1+  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,823  
FILING DATE: April 25, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew Jacob  
REGISTRATION NUMBER: 25,154  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acid residues  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: internal fragment  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:

VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-428-823-7  
Query Match: 48.9%; Score 22; DB 1; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.5e+05;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LHDLL 7  
DB 2 VHDLV 7  
RESULT 13  
US-08-428-823-9  
Sequence 9, Application US/08428823  
Patent No. 5753482  
GENERAL INFORMATION:  
APPLICANT: Yoshizumi ISHINO et al.  
TITLE OF INVENTION: METHOD FOR CLONING OF A GENE FOR  
POLYMERASE  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1+  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,823  
FILING DATE: April 25, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew Jacob  
REGISTRATION NUMBER: 25,154  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acid residues  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: internal fragment  
ORIGINAL SOURCE:  
ORGANISM: Thermus aquaticus  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:

```
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-08-428-823-9
```

```
Query Match 48.9% Score 22; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 LHDDL 7
DB 2 VHDELV 7
```

```
RESULT 14
US-08-556-419-14
; Sequence 14, Application US/08556419C
; Patent No. 6093549
; GENERAL INFORMATION:
; APPLICANT: Ross Christopher
; APPLICANT: Li, Xiao-Jiang
; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan
; APPLICANT: Lanahan, Anthony
; APPLICANT: Worley, Paul
; APPLICANT: Snyder, Solomon
; TITLE OF INVENTION: Huntington-associated protein
; FILE REFERENCE: 0107.52271
; CURRENT APPLICATION NUMBER: US/08/556,419C
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-08-556-419-14
```

```
Query Match 48.9% Score 22; DB 3; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 DDLE 8
DB 1 DDLEQ 5
```

```
RESULT 15
US-09-173-941-82
; Sequence 82, Application US/09173941
```

```
; Patent No. 6140081
; GENERAL INFORMATION:
; APPLICANT: BARBAS, Carlos F.
; TITLE OF INVENTION: ZINC FINGER BINDING DOMAINS FOR GNN
; FILE REFERENCE: NOV0081S
; CURRENT APPLICATION NUMBER: US/09/173,941
; CURRENT FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 82
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: nucleotide
; OTHER INFORMATION: codon binding sequence
; US-09-173-941-82
```

```
Query Match 48.9% Score 22; DB 3; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 3; Gaps 0;
```

```
QY 3 HDDLL 7
DB 2 HDSLL 6
```

```
Search completed: October 20, 2003, 13:50:06
Job time : 23 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:42:21 ; Search time 24 Seconds  
(without alignments)  
36.063 Million cell updates/sec

Title: SEQ15

Perfect score: 42

Sequence: 1 vlrdllea 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168662 residues

Total number of hits satisfying chosen parameters: 789

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 761\*

2: PIR1\*

3: PIR3\*

4: PIR4\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	38.1	7	2	hucolin, 75K chain
2	16	38.1	8	2	hucolin, 75K chain
3	14	33.3	7	2	hucolin, 75K chain
4	14	33.3	7	2	hucolin, 75K chain
5	13	31.0	9	2	hucolin, 75K chain
6	12	28.6	4	2	hucolin, 75K chain
7	12	28.6	5	2	hucolin, 75K chain
8	12	28.6	5	2	hucolin, 75K chain
9	12	28.6	6	2	hucolin, 75K chain
10	12	28.6	6	2	hucolin, 75K chain
11	12	28.6	7	2	hucolin, 75K chain
12	12	28.6	7	2	hucolin, 75K chain
13	12	28.6	7	2	hucolin, 75K chain
14	12	28.6	7	2	hucolin, 75K chain
15	12	28.6	7	2	hucolin, 75K chain
16	12	28.6	8	2	hucolin, 75K chain
17	12	28.6	8	2	hucolin, 75K chain
18	12	28.6	9	2	hucolin, 75K chain
19	12	28.6	9	2	hucolin, 75K chain
20	12	28.6	9	2	hucolin, 75K chain
21	12	28.6	9	2	hucolin, 75K chain
22	12	28.6	9	2	hucolin, 75K chain
23	11	26.2	5	2	hucolin, 75K chain
24	11	26.2	5	2	hucolin, 75K chain
25	11	26.2	5	2	hucolin, 75K chain
26	11	26.2	7	2	hucolin, 75K chain
27	11	26.2	8	2	hucolin, 75K chain
28	11	26.2	8	2	hucolin, 75K chain
29	11	26.2	8	2	hucolin, 75K chain

30 11 26.2 9 2 B28495  
31 11 26.2 9 2 B28495  
32 11 26.2 9 2 B28495  
33 11 26.2 9 2 B28495  
34 11 26.2 9 2 B28495  
35 11 26.2 9 2 B28495  
36 11 26.2 9 2 B28495  
37 11 26.2 9 2 B28495  
38 11 26.2 9 2 B28495  
39 10 23.8 4 2 A48360  
40 10 23.8 4 2 A48360  
41 10 23.8 4 2 A48360  
42 10 23.8 4 2 A48360  
43 10 23.8 6 2 E48394  
44 10 23.8 7 2 E48394  
45 10 23.8 7 2 E48394

#### ALIGNMENTS

##### RESULT 1

S68004  
hucolin, 75K chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S68004  
R:EDGAR, F.F.  
FEBS Lett. 175, 159-161, 1995  
A:Title: hucolin, a new corticosteroid-binding protein from human plasma with structural similarity to the glucocorticoid receptor  
A:Reference number: S68004; MUID:96087107; PMID:7498469  
A:Accession: S68004  
A:Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <EDG>

Query Match 38.1%; Score 16; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6

DB 4 DDL 6

##### RESULT 2

PC4131  
hucolin, 75K chain - human (fragment)  
C:Species: Pseudomonas aeruginosa  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: PC4131  
R:KAWASAKI, S.; ARAI, H.; IGARASHI, Y.; KODAMA, T.  
Gene 167, 87-94, 1995  
A:Title: Sequencing and characterization of the downstream region of the genes encoding hucolin and hucolin-like protein from Pseudomonas aeruginosa  
A:Reference number: JC4552; MUID:96144254; PMID:8566817  
A:Accession: PC4131  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-8 <RAW>  
A:Cross-references: DDBJ:D50473; MUID:G1217594  
A:Note: This ORF is not annotated in GenBank entry PSENIARC, release 113.0

Query Match 38.1%; Score 16; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6

DB 2 DDL 4

##### RESULT 3

```

RESULT 6
I40697
  biotin A - Citrobacter freundii (fragment)
  CSpecies: Citrobacter freundii
  CDate: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-1996
  CAccession: I40697
  R:Shivan, D.; Campbell, A.
  Gene 67, 203-211, 1988
  A>Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter
  A:Reference number: I40697; MUID:89060280; PMID:2971595
  A:Accession: I40697
  A>Status: preliminary; translated from: GB/EVBL/CDRC
  A:Molecule type: DNA
  A:Residues: 1-4 <RES>
  A:Cross-references: GB:X21922; NID:G144434

Query Match      28.6%; Score 12; DB 2; Length 4;
Best Local Similarity 100.0%; Pref. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      4 DD 5
        ||
DB      3 DD 4

RESULT 7
PT601
  T-cell receptor beta chain V-D-J region (120-2K) - mouse (fragment)
  CSpecies: Mus musculus (house mouse)
  CDate: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
  CAccession: PT601; PT0617; PT0694
  R:Feeney, A.J.
  J. Exp. Med. 174, 115-124, 1991
  A>Title: Functional sequences of fetal T cell receptor beta chains have few N regions.
  A:Reference number: PT0509; MUID:91277601; PMID:171158
  A:Accession: PT601
  A>Status: translation: not shown

```

```

A:Residues: 1-5 <FE3>
A:Experimented: source: newborn thymus, strain BALB/c, clone 120-2K
A:Accession: F06617
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FE3>
A:Experimented: source: newborn thymus, strain BALB/c, 120-2CA
A:Accession: F06894
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FE3>
A:Experimented: source: day 18 fetal thymus, strain BALB/c, 154-1H
C:Keywords: T-cell receptor

Query Match      28.6%; Score 12; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches      2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cq      4 DD 5
      ||
Db      4 DD 5

RESULT 8
PT0679

```

C: Carc receptor beta chain v3 region;  
C: Species: Mus musculus (house mouse)  
C: Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C: Accession: P00679; P00678  
R: Peeney, A. J.  
J: Exp. Med. 174, 115-124, 1993.  
A: Title: Functional sequences of fetal T cell receptor beta chains have few N regions.  
A: Reference number: P0039; M01D:9127760.; PMID:171159  
A: Accession: P00679

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-5 <FE>

A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-25

A:Accession: PTC0708

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-5 <FE>

A:Experimental source: newborn thymus, strain: BALB/c, 161-25

C:Keywords: T-cell receptor

Query Match 28.6%; Score 12; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DD 5

Db 4 DD 5

RESULT 9

B35640

Cerebellar degeneration-related protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 28-Sep-1990 #sequence\_revision 28-Sep-1990 #text\_change 24-Jun-1993

C:Accession: B35640

R:Chen, Y.T.; Rettig, W.J.; Yenamandra, A.K.; Kozak, C.A.; Chaganti, R.S.K.; Posner, J.B.

Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990

A>Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal marker

A:Reference number: A35640; MUID:90222173; PMID:2326268

A:Accession: B35640

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-6 <CHE>

Query Match 28.6%; Score 12; DB 2; Length 6;

Best Local Similarity 66.7%; Pred. No. 2.9e+05;

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DD 6

Db 3 DD 5

RESULT 10

PTC533

T-cell receptor beta chain V-D-J region (126-1AA) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PTC533

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A>Title: Functional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PTC533; MUID:91277601; PMID:1711558

A:Accession: PTC533

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-6 <FE>

A:Experimental source: adult thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 12; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DD 5

Db 4 DD 5

RESULT 11

A34026

acetylcholinesterase (EC 3.1.1.7) 5.6S form - Pacific electric ray (fragment)

C:Species: Torpedo californica (Pacific electric ray)

C>Date: 16-Mar-1990 #sequence\_revision 16-Mar-1990 #text\_change 08-Nov-1996

C:Accession: A34026

R:Gibney, G.; MacPhee-Quigley, K.; Thompson, B.; Vedvick, T.; Low, M.G.; Taylor, S.S.;

J. Biol. Chem. 263, 1140-1145, 1988

A>Title: Divergence in primary structure between the molecular forms of acetylcholines

A:Reference number: A34026; MUID:88087239; PMID:3335534

A:Accession: A34026

A:Molecule type: protein

A:Residues: 1-7 <QIB>

C:Keywords: alternative splicing; carboxylic ester hydrolase

Query Match 28.6%; Score 12; DB 2; Length 7;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LLEA 9

Db 1 LUNA 4

RESULT 12

B39043

Calsequestrin, fast skeletal muscle - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 23-Feb-1997

C:Accession: B39040

R:Caia, S.E.; Jones, L.R.

J. Biol. Chem. 266, 391-398, 1991

A>Title: Phosphorylation of cardiac and skeletal muscle calsequestrin isoforms by case

A:Reference number: A39040; MUID:91093153; PMID:1985907

A:Accession: B39040

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-7 <CAL>

C:Keywords: phosphoprotein; skeletal muscle

Query Match 28.6%; Score 12; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DD 5

Db 1 DD 2

RESULT 13

PTC628

T-cell receptor beta chain V-D-J region (111-1AG) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PTC628

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PTC628; MUID:91277601; PMID:1711558

A:Accession: PTC628

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-7 <FE>

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 12; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DD 5

Db 4 DD 5

RESULT 14

PT0722  
T-cell receptor beta chain V-D-J region (135-1G) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0722  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0722  
A>Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-7 <PEE>  
A:Experimental source: newborn thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 28.6%; Score 12; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred.No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5  
||  
DB 4 DD 5

RESULT 15  
PT0576  
T-cell receptor beta chain V-D-J region (141-1G) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0576  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0576  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-7 <PEE>  
A:Experimental source: day 19 fetal thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 28.6%; Score 12; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred.No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5  
||  
DB 4 DD 5

Search completed: October 20, 2003, 13:49:19  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: October 20, 2003, 13:32:06 ; Search time 13 seconds  
(without alignments)

32.557 Million cell updates/sec

Title: SEQ15

Perfect score: 42

Sequence: 1 vlrdlilea 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47C26705 residues

Total number of hits satisfying chosen parameters: 251

Minimum DB seq length: 3

Maximum DB seq length: 9

Post-processing: Maximum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : SwissProt\_41\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	31.0	9	1	IPYR_RHOVI
2	12	28.6	5	1	BIOA_CITFR
3	12	28.6	9	1	NSK1_SARBU
4	12	28.6	9	1	RE42_LITRU
5	12	28.6	9	1	UF02_MOUSE
6	12	28.6	9	1	UFA2_HUMAN
7	11	26.2	8	1	UC26_MAIZE
8	11	26.2	9	1	CONO_CONST
9	11	26.2	9	1	PANRE
10	11	26.2	9	1	PANRE
11	11	26.2	9	1	PANRE
12	11	26.2	9	1	PANRE
13	10	23.8	3	1	LUXE_VIBFI
14	10	23.8	4	1	FLRN_ATEFL
15	10	23.8	6	1	TRPI_PSEPU
16	10	23.8	7	1	CARP_WYTED
17	10	23.8	7	1	PH2_LYCES
18	10	23.8	8	1	MP1_PERAT
19	10	23.8	9	1	LMIP_LOEMI
20	10	23.8	9	1	ULAH_HUMAN
21	9	21.4	4	1	PANRE
22	9	21.4	4	1	PANRE
23	9	21.4	4	1	PANRE
24	9	21.4	7	1	ALJ2_CARNA
25	9	21.4	7	1	PANRE
26	9	21.4	7	1	PANRE
27	9	21.4	7	1	PANRE
28	9	21.4	7	1	PANRE
29	9	21.4	7	1	PANRE
30	9	21.4	7	1	PANRE
31	9	21.4	7	1	PANRE
32	9	21.4	8	1	PANRE
33	9	21.4	8	1	PANRE

34	9	21.4	8	1	FAR2_MACRS
35	9	21.4	8	1	FAR3_HOMAM
36	9	21.4	8	1	FAR4_HOMAM
37	9	21.4	8	1	FAR4_MACRS
38	9	21.4	8	1	FAR8_CALVO
39	9	21.4	8	1	PLP_BRANA
40	9	21.4	9	1	FAR1_CALVO
41	9	21.4	9	1	FAR2_PANRE
42	9	21.4	9	1	FAR3_MACRS
43	9	21.4	9	1	FAR3_PENMO
44	9	21.4	9	1	FAR4_PENMO
45	9	21.4	9	1	FAR5_PENMO

ALIGNMENTS

RESULT 1

IPYR\_RHOVI STANDARD; PRT; 9 AA.

AC P62952

CT 16-OCT-2001 (Rel. 40, Created)

ET 16-OCT-2001 (Rel. 40, Last sequence update)

ET 28-FEB-2003 (Rel. 41, Last annotation update)

DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-

DE hydrolase) (Pase) (Fragment).

GN ZPA

OS Rhodospirillum rubrum

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Hyphomicrobiales; Elstonchloris.

OX NCBI\_TaxID=1079;

RN [1]

RP SEQUENCE, FUNCTION, AND SUBCELLULAR LOCATION.

RA Gomez R., Losada M., Serrano A.;

RE Submitted (JUN-2001) to the SWISS-PROT data bank.

CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS PROTEIN IS:

CC 23 kDa.

CC -!- SIMILARITY: Belongs to the Pase family.

DR HAMAP; MF 00205; ; 1.

DR InterPro; IPR001596; Pyrophosphatase.

DR PROSITE; PS00387; PPASE; PARTIAL.

KW Hydrolase.

FT NON-TER.

SQ SEQUENCE 9 AA; 1014 MW; 68F3EAA05DDAA044 CRC64;

Query Match: 31.0%; Score 13; DB 1; Length 9;

Best Local Similarity 28.6%; Pred. No. 1.3e+05;

Matches 2; Conservative 3; Mismatches 2; Indels 0;

Gaps 0;

QY 2 LRDULE 9

DE 1 KRDAID 7

RESULT 2

BIOA\_CITFR STANDARD; PRT; 5 AA.

AC P13071

CT 21-JAN-1990 (Rel. 13, Created)

ET 21-JAN-1990 (Rel. 13, Last sequence update)

ET 16-OCT-2001 (Rel. 40, Last annotation update)

DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase

DE (EC 2.6.1.62) (7,8-diamino-9-oxononanoic acid aminotransferase) (DAPA

DE aminotransferase) (Fragment).

GN BCOA

OS Citrobacter freundii.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Citrobacter.

OX NCBI\_TaxID=546;

RN [1]

RP SEQUENCE FROM N.A.





RT "Separation and sequencing of familiar and novel marine proteins using preparative two-dimensional gel electrophoresis.";  
 RL Electrophoresis 15:733-745(1994).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6.0, ITS MW IS: 32 kDa.  
 FT NON TER 9  
 SQ SEQUENCE 9 AA; 1102 MW; 7E73EAB6D05B1AA5 CRC64;

Query Match 26.6%; Score 12; DB 1; Length 9;  
 Best Local Similarity 20.3%; Pred. No. 1.3e+05;  
 Matches 1; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LRDDL 6  
 DB 1 MEDEI 5

RESULT 6  
 UHA2 HUMAN STANDARD; PRT; 9 AA.  
 AC P40929;  
 DT 01-FEB-1995 (Rel. 31, Last sequence update);  
 DT 01-FEB-1995 (Rel. 31, Last sequence update);  
 DT 16-OCT-2001 (Rel. 40, Last annotation update);  
 DE Unknown protein from 2D-page of heart (Spot 5603) (Fragment).  
 CS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.X., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.C.;  
 RT "The human myocardin two-dimensional gel protein database: update 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6.0, ITS MW IS: 55.3 kDa.  
 FT NON TER 9  
 SQ SEQUENCE 9 AA; 1104 MW; 8874B1BB5B01B2CA CRC64;

Query Match 28.6%; Score 12; DB 1; Length 9;  
 Best Local Similarity 100.3%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 DD 5  
 DB 6 DD 5

RESULT 7  
 UC26 MAIZE STANDARD; PRT; 8 AA.  
 AC P80632;  
 DT 01-OCT-1996 (Rel. 34, Created);  
 DT 01-OCT-1996 (Rel. 34, Last sequence update);  
 DT 15-JUL-1999 (Rel. 38, Last annotation update);  
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 907; (Fragment)).  
 CS Zea mays (Maize).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC PACAD clade; Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE=Coleoptile;  
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
 RA Pernollet J.-C., Zivy M., de Vienne D.;  
 RT "The maize two dimensional gel protein database: towards an integrated genome analysis program.";  
 RL Theor. Appl. Genet. 93:997-1005(1996).

CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 7.0, ITS MW IS: 57.2 kDa.  
 ER Maize-2DPAGE; P80632; COLEOPTILE.  
 FT NON TER 1  
 SQ SEQUENCE 8 AA; 990 MW; 9639D6DAB4176B1D CRC64;

Query Match 26.2%; Score 11; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RD 4  
 DB 4 RD 5

RESULT 8  
 CCNO CONST STANDARD; PRT; 9 AA.  
 AC P05487;  
 DT 01-NOV-1988 (Rel. 39, Created);  
 DT 01-NOV-1988 (Rel. 39, Last sequence update);  
 DT 28-FEB-2003 (Rel. 41, Last annotation update);  
 DE Arg-conopressin S.  
 CS Conus striatus (Striated cone).  
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda; Apogastropoda; Caenogastropoda; Scrobocioncha; Hypogastropoda; Negastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=6493;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=88058932; PubMed=3680228;  
 RA Cruz L.J., de Santos V., Zafaralla G.C., Ramilo C.A., Zeikus R.D.,  
 RA Gray W.R., Oliveira B.M., Cruz L.C.;  
 RT "Peptide toxins from venomous Conus snails.";  
 RL Ann. Rev. Biochem. 57:665-700(1988).  
 CC -1- FUNCTION: Targets vasopressin-oxytocin related receptors.  
 CC -1- SUBCELLULAR LOCATION: Secreted by the venom duct.  
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.  
 CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR PIR; B28495; B28495.  
 DR InterPro; IPR000991; Neurhyp\_horm.  
 DR Pfam; PF00220; hormone4; 1.  
 DR PROSITE; PS00264; NEUROHYPOPHYS\_HORM; 1.  
 KW Hormone; Amidation.  
 FI DSJLFD  
 FT MOD\_RES 9  
 SQ SEQUENCE 9 AA; 1831 MW; 17EB76EB4540050 CRC64;

Query Match 26.2%; Score 11; DB 1; Length 9;  
 Best Local Similarity 25.0%; Pred. No. 1.3e+05;  
 Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLRD 4  
 DB 2 LRN 5

RESULT 9  
 PARS PANRE STANDARD; PRT; 9 AA.  
 AC P82861;  
 DT 16-OCT-2001 (Rel. 40, Created);  
 DT 16-OCT-2001 (Rel. 40, Last sequence update);  
 DT 16-OCT-2001 (Rel. 40, Last annotation update);  
 DE PVRamide-like neuropeptide PFS (AMRNALVRF-amide).

DS Panagrellus redivivus.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.  
OC NCBI\_TaxID=6233;  
CX [1]  
CX SEQUENCE, FUNCTION, AND AMIDATION.  
RP Moffet C.L., Marks N.J., Halton D.W., Thomson D.P., Geary T.G.,  
RA Maule A.G.;  
RA "Isolation, characterization and pharmacology of FMRFamide-related  
RT peptides (FaRPs) from free-living nematode, Panagrellus redivivus";  
RT Submitted (JUL-2000) to the SWISS-PROT data bank.  
CC -!- FUNCTION: MWACTIVE.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
CC Neuropeptide; Amidation.  
CC MOD RES 9  
CC SEQUENCE 9 AA; 1077 MW; A0D112C72DD45406 CRC64;  
Query Match 26.2%; Score 11; DB 1; Length 9;  
Best Local Similarity 33.3%; Pred. No. 1.3e+05;  
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Y 2 LRDBLL 7  
DB 2 MRNALV 7  
RESULT 10  
FARA CALVO STANDARD; PRT; 9 AA.  
ID FARA CALVO  
AC P4.865;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CalliFMRFamide 10.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
OC Calliphoridae; Calliphora.  
OC NCBI\_TaxID=27454;  
RN [1]  
RP SEQUENCE.  
RP TISSUE=Thoracic ganglion; PubMed=1549595;  
RX MEDLINE=92196111; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
RA Rehfeld J.F., Thorpe A.;  
RA "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
RT neuropeptides (designated calliFMRFamides) from the blowfly  
RT Calliphora vomitoria";  
RT Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
RL -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
CC PIR: A44787;  
DR Neuropeptide; Amidation.  
KW MOD\_RES 9  
KW UNSURE 1 1  
KW OR S OR A.  
FT SEQUENCE 9 AA; 1183 MW; 29D0C699CAB40457 CRC64;  
SQ SEQUENCE 9 AA; 1183 MW; 29D0C699CAB40457 CRC64;  
Query Match 26.2%; Score 11; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 3 RD 4  
DB 4 RD 5  
RESULT 11  
PTSP BOMVO  
ID PTSP BOMVO STANDARD; PRT; 9 AA.  
AC P82003;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Prothoracicostatic peptide (Bom-PTSP).  
OC Bombyx mori (Silk moth).  
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Bombycoidea;  
OC Bombycidae; Bombyx.  
CX NCBI\_TaxID=7091;  
CX [1]  
CX SEQUENCE.  
RP STRA:NC=45 X N140; TISSUE=Brain;  
RX MEDLINE=20026334; PubMed=10531308;  
RA Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara K., Nagata S.,  
RA Kataoka H.;  
R: "Identification of a prothoracicostatic peptide in the larval brain: of  
RT the silkworm, Bombyx mori";  
RL J. Biol. Chem. 274:31169-31172(1999).  
RN [2]  
RN PERATUM.  
RP Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara K., Nagata S.,  
RA Kataoka H.;  
RL J. Biol. Chem. 275:9892-9892(2000).  
CC -!- FUNCTION: Inhibits ecdysteroid biosynthesis in the prothoracic  
CC gland.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DEVELOPMENTAL STAGE: EARLY FIFTH INSTAR.  
KW Hormone; Amidation.  
KW MOD\_RES 9  
KW SEQUENCE 9 AA; 1090 MW; 3878C594472AB6C3 CRC64;  
Query Match 26.2%; Score 11; DB 1; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Y 5 DLLEA 9  
DB 4 DLNSA 9  
RESULT 12  
TALI PICJA STANDARD; PRT; 9 AA.  
ID TALI PICJA  
AC P17470;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transaldolase 1 (EC 2.2.1.2) (Fragment).  
OS Pichia jadinii (Yeast) (Candida utilis).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Pichia.  
CX NCBI\_TaxID=4903;  
CX [1]  
CX SEQUENCE.  
RX MEDLINE=77110646; PubMed=556924;  
RA Sun S.C., Joris J., Tsoias O.;  
RT "Purification of crystallization of transaldolase isozyme 1 and  
RT evidence for different genetic origin of isozymes 1 and 2 in  
RT Candida utilis";  
RL Arch. Biochem. Biophys. 178:69-78(1977).  
CC -!- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF  
CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.  
CC -!- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde  
CC 3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.  
CC -!- PATHWAY: Pentose phosphate pathway; nonoxidative part.  
CC -!- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY. SUBFAMILY 1.  
DR InterPro: IPR001585; Transaldolase.  
DR PIR: A12872; A12872.  
DR PROSITE: PS00358; TRANSALDOLASE\_2; PARTIAL.  
DR PROSITE: PS01354; TRANSALDOLASE\_1; PARTIAL.  
KW Transferase; Pentose shunt.  
KW NON\_TER 1  
KW NON\_TER 9  
FT SEQUENCE 9 AA; 1008 MW; 274F31AF0E51E059 CRC64;  
SQ SEQUENCE 9 AA; 1008 MW; 274F31AF0E51E059 CRC64;

```

Query Match          26.2%; Score 11; DB 1; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 DBLL 7
DB 6 BTL 9

RESULT 13
LUXE_VIBFI
ID LUXE_VIBFI STANDARD; PRT; 3 AA.
AC P24272;
DT 21-MAR-1992 (Rel. 21; Created);
DT 01-MAR-1992 (Rel. 21; Last sequence update);
DT 01-OCT-1996 (Rel. 34; Last annotation update);
DE Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl-
protein synthetase) (Fragment);
GN LUXE;
OS Vibrio fischeri;
CC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
CC Vibrionaceae; Vibrio.
CX NCBI_TaxID=663;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=9:072226; PubMed=2254256;
RA Swartzman E., Kapoor S., Graham A.F., Meighen F.A.;
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination
site for the lux operon.";
RL J. Bacteriol. 172:6797-6802(1990).
CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
CC -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
CC an acyl-protein thioester.
CC -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

SEQUENCE, AND MASS SPECTROMETRY.
RX MEDLINE=9:319122; PubMed=1973541;
RA Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
RA Reinscheid R.K., Nockher H.-P., Staley A.L.;
RT "Isolation of L-3-phenyllactyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea
anemone neuropeptide containing an unusual amino-terminal blocking
group.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron-specific.
CC -!- MASS SPECTROMETRY: MW=543.3; METHOD=FAB.
DR PIR; A35779; A35779.
KW Neuropeptide; Amidation.
FT VCD RES 1 L-3-PHENYLLACTYL.
FT MOD_RES 4 4 AMIDATION
SQ SEQUENCE 4 AA; 549 MW; 64540729AC00000000 CRC64;

Query Match          23.8%; Score 10; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LRD 4
DB 2 LRN 4

RESULT 15
TRPI_PSEPU
ID TRPI_PSEPU STANDARD; PRT; 6 AA.
AC P36414;
DT 01-JUN-1994 (Rel. 29; Created);
DT 01-JUN-1994 (Rel. 29; Last sequence update);
DT 01-FEB-1995 (Rel. 31; Last annotation update);
DE TRPA operon: transcriptional activator (Fragment).
GN TRPI.
OS Pseudomonas putida.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
CC Pseudomonadaceae; Pseudomonas.
CX NCBI_TaxID=303;
RN [1];
RP SEQUENCE FROM N.A.
RX STRAIN=PPG1 C15;
RX MEDLINE=89335826; PubMed=2503057;
RA Eberly L., Crawford I.P.;
RT "DNA sequence of the tryptophan synthase genes of Pseudomonas
putida.";
RL Biochimie 71:521-531(1989).
CC -!- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPA GENES ENCODING
CC THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE
CC INDUCER (INDOLEGLYCEROL PHOSPHATE), TRP1 BINDS UPSTREAM OF THE
CC TRPA OPERON, OVERLAPPING ITS OWN PROMOTER REGION.
CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

Qy : 5 DJ 6  
    ||  
Db : 4 DL 5

Search completed: October 20, 2003, 13:46:19  
Job time : 13 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode:

Run on: October 20, 2003, 13:41:46 ; Search time 59 Seconds  
(without alignments)  
39.364 Million cell updates/sec

Title: SEQ1B  
Perfect score: 42  
Sequence: 1 virddllea 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 773

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTRXBL 23:\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacterioph:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	38.1	9	2 P82568	P82568 streptococ
2	16	39.1	9	10 Q8LPT5	Q8LPT5 zea mays im
3	15	35.7	8	2 P72279	P72279 rhodococcus
4	15	35.7	9	4 Q9UE26	Q9UE26 homo sapien
5	15	35.7	9	9 Q9XJ40	Q9XJ40 bacterioph
6	14	33.3	8	2 Q8KPX4	Q8KPX4 microcystis
7	14	33.3	8	2 Q9Z159	Q9Z159 neisseria m
8	14	33.3	8	2 Q51594	Q51594 escherichia
9	14	33.3	9	13 Q9FSZ2	Q9FSZ2 cicer ariet
10	14	33.3	9	11 Q9QZ48	Q9QZ48 mus musculu
11	14	33.3	9	15 Q85710	Q85710 rous sarcom
12	14	33.3	9	15 Q8UTD7	Q8UTD7 human immun
13	13	31.0	7	4 Q15857	Q15857 hmo sapien
14	13	31.0	8	2 Q93SR0	Q93SR0 staphylococ
15	13	31.0	8	3 Q9HDS4	Q9HDS4 aspergillus
16	13	31.0	8	11 Q9QWJ8	Q9QWJ8 mus sp. mep

17	13	31.0	8	12 Q89965	Q89965 polyomaviru
18	13	31.0	9	12 Q91BM8	Q91BM8 simiar viru
19	13	31.0	9	12 Q9PYK1	Q9PYK1 simiar viru
20	12	28.6	7	8 Q99182	Q99182 gnatholebia
21	12	28.6	7	10 Q49223	Q49223 G-lycine max
22	12	28.6	8	5 Q9NGM5	Q9NGM5 toxoplasma
23	12	28.6	8	6 Q8WNS1	Q8WNS1 bos taurus
24	12	28.6	8	8 Q35792	Q35792 saccharomyc
25	12	28.6	8	9 Q8SBU0	Q8SBU0 bacterioph
26	12	28.6	8	9 Q8H9K4	Q8H9K4 bacterioph
27	12	28.6	8	9 Q8H9K1	Q8H9K1 bacterioph
28	12	28.6	8	9 Q8H9J9	Q8H9J9 bacterioph
29	12	28.6	8	9 Q8H9J7	Q8H9J7 bacterioph
30	12	28.6	8	9 Q8H9J5	Q8H9J5 bacterioph
31	12	28.6	8	9 Q8H9J3	Q8H9J3 bacterioph
32	12	28.6	8	9 Q8H9J1	Q8H9J1 bacterioph
33	12	28.6	8	9 Q8H9I9	Q8H9I9 bacterioph
34	12	28.6	8	9 Q8H9I8	Q8H9I8 bacterioph
35	12	28.6	8	9 Q8H9I6	Q8H9I6 bacterioph
36	12	28.6	8	9 Q8H9I4	Q8H9I4 bacterioph
37	12	28.6	8	9 Q8H9I3	Q8H9I3 bacterioph
38	12	28.6	8	9 Q8H9I2	Q8H9I2 bacterioph
39	12	28.6	8	9 Q8H9H6	Q8H9H6 bacterioph
40	12	28.6	8	9 Q8H9H5	Q8H9H5 bacterioph
41	12	28.6	8	9 Q8H9H3	Q8H9H3 bacterioph
42	12	28.6	8	13 Q9S669	Q9S669 gallus gall
43	12	28.6	8	15 Q98YK9	Q98YK9 human immun
44	12	28.6	9	4 Q9BQT4	Q9BQT4 hmo sapien
45	12	28.6	9	13 Q8AYJ5	Q8AYJ5 carassius a

## ALIGNMENTS

## RESULT 1

P82568 PRELIMINARY; PRT; 9 AA.  
AC P82568  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE Unknown protein from 2D-page (Fragment).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_taxid=1314;  
RN [1]  
RP SEQUENCE, AND MASS SPECTROMETRY.  
RC STRAIN=JRS4;  
RA Hogan D.A., D.J.P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,  
RA VanBogelen R.A.;  
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes  
proteins";  
RL Submitted (MAY-2000) to the SWISS-PROT data bank.  
CC -!- MASS SPECTROMETRY: MW=22592.04; METHOD=ELECTROSPRAY.  
FT NCN TER 1 1  
FT NCN TER 9 9  
SQ SEQUENCE 9 AA; 1369 MW; 2A771042CB1AB2D7 CRC64;

Query Match 38.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 40.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CY 4 DDLIE 8  
DB 4 DEVIE 8

## RESULT 2

Q8LPT5 PRELIMINARY; PRT; 9 AA.  
ID Q8LPT5  
AC Q8LPT5  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Beta-expansin-like protein (Fragment).  
 OS Zea mays (Maize)  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 CX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. Cl23;  
 RA Ching A.S., Cardwell K.S., Jung X., Dolan M., Smith O.S., Tingey S.,  
 RA Morgante M., Rafalski J.A.;  
 RA "SNP frequency, haplotype structure and linkage disequilibrium in  
 RI elite maize inbred lines."  
 RI Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY094310; AM21936.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 9 AA; 977 MW; 50552D2CB1AAA3 CRC64;

Query Match 38.1%; Score 16; DB 10; Length 9;  
 Best Local Similarity 33.3%; Pred. No. 8.3e+05;  
 Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

CY 4 DDLLEA 9  
 DB 4 DEVVDA 9

## RESULT 3

ID P72279 PRELIMINARY; PRT; 8 AA.  
 AC P72279;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Biphenyl dioxygenase (fragment).  
 GN BPBH.  
 OS Rhodococcus globerulus.  
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 CC Corynebacterineae; Nocardiaceae; Rhodococcus.  
 CX NCBI\_TaxID=33008;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=95255652; PubMed=737502;  
 RA Asturias J.A., Diaz E., Timmis K.N.;  
 RA "Evolutionary relationship of the biphenyl dioxygenase of the gram-  
 RT positive bacterium Rhodococcus globerulus p6 to multicomponent  
 RT dioxygenases of gram-negative bacteria."  
 RL Gene 156:11-18(1995).  
 DR EMBL; X80041; CAA56350.1; -.  
 FT DIOXYGENASE.  
 KW NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 989 MW; EBD2CB1AB6D73406 CRC64;

Query Match 35.7%; Score 15; DB 2; Length 8;  
 Best Local Similarity 33.3%; Pred. No. 8.3e+05;  
 Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

CY 2 LRDLLE 7  
 DB 3 LQDEVV 8

## RESULT 4

ID Q9UE26 PRELIMINARY; PRT; 9 AA.  
 AC Q9UE26;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE PEG1/WEST protein.  
 GN PEG1/WEST.

OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=97336048; PubMed=5192843;  
 RA Riesewijk A.M., Hu L., Schulz U., Tariverdian G., Hoeglund P.,  
 RA Kere J., Rogers H.H., Kalscheuer V.M.;  
 RA "Monoclonal expression of human PEG1/WEST is paralleled by parent-  
 RC specific methylation in fetuses."  
 RL Genomics 42:236-244(1997).  
 DR EMBL; Y15626; CAA1631.1; -.  
 SQ SEQUENCE 9 AA; 1258 MW; C52DC7340AB4412 CRC64;

Query Match 35.7%; Score 15; DB 4; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 8.3e+05;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 VLRDUL 6  
 DB 2 VRDRL 7

## RESULT 5

ID Q9XKNC PRELIMINARY; PRT; 9 AA.  
 AC Q9XKNC;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE P19 (Fragment).  
 OS Bacteriophage phi-10.  
 CC Viruses; dsDNA viruses; Cystoviridae; Cystovirus.  
 CX NCBI\_TaxID=90889;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=99350412; PubMed=10419946;  
 RA Mindich L., Qiao X., Qiao C., Onodera S., Romantschuk M.,  
 RA Hoogstraaten D.;  
 RA "Isolation of additional bacteriophages with genomes of segmented  
 RT double-stranded RNA."  
 RL J. Bacteriol. 181:4505-4508(1999).  
 DR EMBL; AF125675; AAC22555.1; -.  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1058 MW; 880E376AA720544A CRC64;

Query Match 35.7%; Score 15; DB 9; Length 9;  
 Best Local Similarity 40.8%; Pred. No. 8.3e+05;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CY 4 DDLLE 8  
 DB 2 DNLID 6

## RESULT 6

ID Q8KPY4 PRELIMINARY; PRT; 8 AA.  
 AC Q8KPY4;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Phycocyanin alpha subunit (fragment).  
 GN PCA.

OS Microcystis sp. T95-1.  
 CC Bacteria; Cyanobacteria; Chroococcales; Microcystis.  
 CX NCBI\_TaxID=198099;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STAN=biom water sample T96-1;  
 RA Baker J.A., Entsch B., Neilan B.A., McKay D.B.;  
 RA "Monitoring changing toxigenicity of a cyanobacterial bloom using

```

RT Molecular methods.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY117046; AAM54719.1; -.
FT NON_TER 8
SO SEQUENCE 8 AA; 890 MW; F4CB01A73771A336 CRC64;

Query Match 33.3%; Score 14; DB 2; Length 8;
Best Local Similarity 37.5%; Pred. No. 8.3e+05;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LRDDLLEA 9
DB 1 MKTFLTER 8

RESULT 7
ID Q9ZIE9 PRELIMINARY; PRT; 8 AA.
AC Q9ZIE9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update);
DE Carbamoyl-phosphate synthase subunit B (Fragment).
GN CARB.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]_TaxID=487;
RP SEQUENCE FROM N.A.
RX STRAIN=1527;
RA Lawson F.S., Billows F.M., Dillon J.A.;
RT "Organization of carbamoyl-phosphate synthase genes in Neisseria gonorrhoeae includes a large, variable intergenic sequence which is also present in other Neisseria species.";
RL Microbiology 141:0-0(0).
RN [2]_TaxID=487;
RP SEQUENCE FROM N.A.
RX STRAIN=1527;
RA Brinkman F.S.L., Francis F.W., Dillon J.R.;
RT "Complexity of the variable sequence between the carbamoyl-phosphate synthase genes of Neisseria species.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF029361; AAC78449.1; -.
FT NON_TER 8
SO SEQUENCE 8 AA; 988 MW; FAJ72AB134032766 CRC64;

Query Match 33.3%; Score 14; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 9.3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLE 6
DB 4 RTDL 7

RESULT 8
ID Q51594 PRELIMINARY; PRT; 8 AA.
AC Q51594;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update);
DE CopB protein (Fragment).
OS Escherichia coli.
OG Plasmid ColV2-K94.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]_TaxID=562;
RP SEQUENCE FROM N.A.
RX MEDLINE=8623772; PubMed=2423502;

```

```

RA Weber P.C., Palchaudhuri S.;
RT "Incompatibility repressor in a repA-like replicon of the IncPI plasmid ColV2-K94.";
RL E. Bacteriol. 166:1106-1112(1986).
DR EMBL; M13472; AAA21194.1; -.
KW Plasmid.
FT NON_TER 1
SO SEQUENCE 8 AA; 998 MW; 0F37205AA73416D7 CRC64;

Query Match 33.3%; Score 14; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLE 8
DB 3 RDILK 8

RESULT 9
ID Q9FSZ2 PRELIMINARY; PRT; 9 AA.
AC Q9FSZ2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update);
DE Hypothetical 1.0 kDa protein (Fragment).
OS Cicer arietinum (Chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotids 1; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
OX NCBI_TaxID=3327;
RN [1]_TaxID=3327;
RP SEQUENCE FROM N.A.
RX STRAIN=CV. Castellana; TISSUE=Etisolated epicotyl;
RA Dopico B., Jimenez T., Labrador E.;
RT "cDNA clones expressed in etiolated Cicer arietinum epicotyls.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ299089; CAC10216.1; -.
KW Hypothetical protein.
FT NON_TER 1
SO SEQUENCE 9 AA; 990 MW; 9441BDDAA7272EBE CRC64;

Query Match 33.3%; Score 14; DB 10; Length 9;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 LLEA 9
DB 4 LIDA 7

RESULT 10
ID Q9QZAB PRELIMINARY; PRT; 9 AA.
AC Q9QZAB;
DT 01-MAY-2003 (TrEMBLrel. 13, Created)
DT 01-MAY-2003 (TrEMBLrel. 13, Last sequence update);
DE C-type lectin DCL1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RA Gorski K., Huang X., Tseng S.-Y., Rattis F., Pardoll D., Tsuchiya H.;
RT "Dedritic cell regulation of DCL1 mRNA expression.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF192526; AAF04643.1; -.
DR MGD; MG:236650; Dcl1.
KW Lectin.
FT NON_TER 9

```

SQ SEQUENCE 9 AA; 994 MW; 342161AB172EBAB7 CRC64;

Query Match 33.3%; Score 14; DB 11; Length 9;  
Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 DLLE 8  
DB 3 DCLL 6

RESULT 11

ID Q85710 PRELIMINARY; PRT; 9 AA.  
AC Q85710;  
DT 01-NOV-1996 (TrEMBLrel. 0; Created)  
DT 01-NOV-1996 (TrEMBLrel. 0; Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08; Last annotation update)  
DE Pol protein (Fragment).  
OS Rous sarcoma virus.  
OS Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.  
OC NCBI\_TaxID=11886;  
RN [1]\_TaxID=11886;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=64115080; PubMed=6319754;  
RA Lerner T.L., Hanafusa H.;  
RT "DNA sequence of the Bryan high-titer strain of Rous sarcoma virus:  
RT Extent of env deletion and possible genealogical relationship with  
RT other viral strains";  
RL J. Virol. 49:545-556(1984).  
RL EMBL; K03365; AAA42557.1; -.  
FT NON TER 1  
SQ SEQUENCE 9 AA; 949 MW; 94AA14BDD0731AA CRC64;

Query Match 33.3%; Score 14; DB 15; Length 9;  
Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DDLLE 9  
DB 1 EDTLAA 6

RESULT 12

Q8UTD7 PRELIMINARY; PRT; 9 AA.  
ID Q8UTD7  
AC Q8UTD7  
DT 01-MAR-2002 (TrEMBLrel. 20; Created)  
DT 01-MAR-2002 (TrEMBLrel. 20; Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22; Last annotation update)  
DE Vpu protein.  
GN VPJ.  
OS Human immunodeficiency virus 1.  
OS Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OC NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=00BM1471.27;  
RA Novitsky V.A., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,  
RA Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,  
RA Foley B.T., Gaoletwe S., Rybak N., Gaseitsiwe S., Vannberg F.,  
RA Marink R., Lee T.-H., Essex M.;  
RT "HIV-1 subtype C molecular phylogeny: consensus sequence for an AIDS  
RT vaccine design."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AF443091; AAL34712.1; -.  
SQ SEQUENCE 9 AA; 1102 MW; 1885D40B1727244C CRC64;

Query Match 33.3%; Score 14; DB 15; Length 9;  
Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLE 8

DB 3 NELLE 6

RESULT 13

Q15897 PRELIMINARY; PRT; 7 AA.  
AC Q15897;  
DT 01-NOV-1996 (TrEMBLrel. 01; Created)  
DT 01-NOV-1996 (TrEMBLrel. 01; Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19; Last annotation update)  
DE (Clone X56A11A) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Placenta;  
RA Lee C.-C., Yazdani A., Nehrert M., Bailey J., Couch L., Xiong M.,  
RA Coolbaugh M., Chinault C.A., Baldini A., Lindsey E.A., Zhao Z.-Y.,  
RA Caskey C.T.H.;  
RT "Isolation of chromosome specific genes by reciprocal probing of  
RT arrayed cDNAs and cosmid libraries";  
RL Hum. Mol. Genet. 3:10-3(1995).  
RL EMBL; L32077; AAA73867.1; -.  
FT NON TER 1  
SQ SEQUENCE 7 AA; 314 MW; 672BDD3372045B0 CRC64;

Query Match 31.3%; Score 13; DB 4; Length 7;  
Best Local Similarity 33.3%; Pred. No. 8.3e-05;  
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLKDDL 6  
DB 2 LKXEL 7

RESULT 14

Q93S80 PRELIMINARY; PRT; 8 AA.  
AC Q93S80;  
DT 01-DEC-2001 (TrEMBLrel. 19; Created)  
DT 01-DEC-2001 (TrEMBLrel. 19; Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19; Last annotation update)  
DE Beta-lactamase repressor Bial (Fragment).  
GN BJA1.  
OS Staphylococcus epidermidis.  
OC Bacteri; Firmicutes; Bacillales; Staphylococcus.  
OC NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=6;  
RA StChu M.S., Heir E., Sorum H., Hoick A.L.;  
RT "Genetic linkage between quaternary ammonium compound and beta-lactam  
RT resistance in Staphylococci isolated from food."  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AY028779; AA038453.1; -.  
KW Plasmid.  
FT NON TER 1  
SQ SEQUENCE 8 AA; 930 MW; 4E3325B05AA44720 CRC64;

Query Match 31.3%; Score 13; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLKD 4  
DB 1 ILND 4



```

RESULT 15
QHDS4
ID QHDS4 PRELIMINARY; PRF; 8 AA.
AC QHDS4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
ET 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE TPC polyprotein (Fragment).
GN TPC.
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5059;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A55;
RA Geiser D.M., Dorrer J.W., Horn B.W., Taylor J.W.;
RT "The phylogenetics of mycotoxin and sclerotium production in
RT Aspergillus flavus and Aspergillus oryzae."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF261861; AAC:6135..;
KW Polyprotein.
FT NON-TER
SQ SEQUENCE 8 AA; 907 MW; F3B2C72AB5B57DD6 CRC64;

```

```

Query Match: 31.08; Score 13; DB 3; Length 8;
Best Local Similarity 50.08; Pred. NO. 8.3e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 DLES 8
DB 5 DLVD 8

```

```

Search completed: October 20, 2003, 13:48:25
Job time : 59 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:31:21 : Search time 56 Seconds  
(without alignments)  
25,510 Million cell updates/sec

Title: SEQUIB

Perfect score: 42

Sequence: i vlrdallea 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: i107863 seqs, i58726573 residues

Total number of hits satisfying chosen parameters: 179625

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun01.\*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1983.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	9	20	AAW97374
2	42	100.0	9	20	AAW99197
3	37	88.1	9	20	AAW97375
4	37	88.1	9	20	AAW99196
5	36	85.7	9	20	AAW97572
6	36	85.7	9	20	AAW99195
7	27	64.3	9	19	AAW47427
8	25	59.5	9	23	AAU71428
9	24	57.1	7	22	AAW81550

	10	23	54.8	9	23	AAU71211	Human MHC class I
11	23	54.8	9	24	ABR18988	Human cancer-relat	
12	23	54.8	9	24	ABR19174	Human cancer-relat	
13	23	54.8	9	24	ABR19426	Human cancer-relat	
14	23	54.8	9	24	ABR19578	Human cancer-relat	
15	23	54.8	9	24	ABR19780	Human cancer-relat	
16	23	54.8	9	24	ABR19981	Human cancer-relat	
17	22	52.4	9	20	AAW97373	Peptide epitope of	
18	21	50.0	6	21	AA8:2035	Peptide # 2 used i	
19	21	50.0	6	23	AAU80810	Rat Rb-interacting	
20	21	50.0	7	23	ASG79371	CXR zinc-finger h	
21	21	50.0	7	23	ABP48411	zinc finger protei	
22	21	50.0	7	23	ABP48414	zinc finger protei	
23	21	50.0	7	23	ABP48419	zinc finger protei	
24	21	50.0	7	24	ABP96179	zinc finger nucleo	
25	21	50.0	8	18	AA26557	Soluble peptide in	
26	21	50.0	8	21	AA812089	Ad7 cel peptide.	
27	21	50.0	8	22	AA846535	Integrin alphav-be	
28	21	50.0	8	22	AA846568	Erbb2 cell overexp	
29	21	50.0	8	24	ABP93038	Cel motif peptide	
30	21	50.0	9	21	AA812086	Ad40 cel peptide.	
31	21	50.0	9	21	AA812092	H2A-A2 restricted	
32	21	50.0	9	21	AA812092	Human C35 peptide	
33	21	50.0	9	22	AA812092	Human C35 peptide	
34	21	50.0	9	22	AA812092	Human C35 peptide	
35	21	50.0	9	22	ABR12506	Human C35 peptide	
36	21	50.0	9	22	ABR12554	Human C35 peptide	
37	21	50.0	9	22	ABR12579	Human C35 peptide	
38	21	50.0	9	22	ABR12651	Human C35 peptide	
39	21	50.0	9	22	ABR12658	Human C35 peptide	
40	21	50.0	9	22	ABR12737	Human C35 peptide	
41	21	50.0	9	22	ABR12755	Human C35 peptide	
42	21	50.0	9	22	ABR12774	Human C35 peptide	
43	21	50.0	9	22	ABR12808	Human C35 peptide	
44	21	50.0	9	22	ABR12949	Human C35 peptide	
45	21	50.0	9	22	ABR12983	Human C35 peptide	

#### ALIGNMENTS

RESULT 1  
AAW97374  
ID AAW97374 standard; Protein; 9 AA.  
XX  
AC AAW97374;  
XX  
DT 13-MAY-1999 (first entry)  
XX  
DE HA-1 R-allele sequence.  
XX  
KW Intron; minor histocompatibility antigen HA-i; typing allele;  
KW H allele; R allele; polymorphic nucleotide; HA-1 typing;  
KW bone marrow transplant; severe aplastic anaemia; leukaemia;  
KW immune deficiency disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WC8905313-A2.  
XX  
PC 04-FEB-1999.  
XX  
PF 23-JUL-1998; 98WO-EP04928.  
XX  
PR 02-JUN-1998; 98EP-0870125.  
PR 23-JUL-1997; 97EP-0202303.  
XX  
PA (UYLE-) RICKSUNIV LEIDEN.  
XX  
PI Goulmy E;  
XX  
DR W21: 1999-142960/12.  
XX

PT Typing minor histocompatibility antigen HA-1 - by amplifying and  
 PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g.  
 PT detection of genetic aberrances

PS Claim 13; Fig 5; 59pp; English.

XX The present sequence represents part of the minor histocompatibility  
 CC antigen HA-1 R-allele. The specification describes methods for typing  
 CC alleles (preferably the H and R alleles) of the minor histocompatibility  
 CC antigen HA-1 in a sample, which comprise detecting polymorphic  
 CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The  
 CC methods can be used for HA-1 typing for bone marrow transplants, severe  
 CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as  
 CC detection of genetic aberrances. The probes and primers of the invention  
 CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be  
 CC used anti-idiotypic B cells and/or T cells and antibodies.

XX Sequence 9 AA;

Query Match 100.0%; Score 42; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLKDDLEA 9  
 |||||  
 Db 1 VLKDDLEA 9

RESULT 2  
 ID AAW99197  
 AC AAW99197 standard; peptide: 9 AA.

XX AC AAW99197;  
 XX C: 20-MAY-1999 (first entry)  
 XX DE Minor histocompatibility antigen HA-1 T-cell epitope #3.

XX Minor histocompatibility antigen: HA-1; T-cell epitope; immunological;  
 KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;  
 KW diagnosis; aplastic anaemia; immune deficiency disease.

XX Homo sapiens.

XX WO9905174-A1.

XX 04-FEB-1999.

XX 23-JUL-1998; 98WO-NL00425.

XX 23-JUL-1997; 97EP-0202303.

XX (UYLE-); RIJKSUNIV LEIDEN.

XX Engelhard VH, Goulmy EAJM, Hunt DF;

XX WPI; 1999-153312/13.

XX A new minor histocompatibility antigen, HA-1 - useful to treat  
 PT immune diseases and prevent rejection and host versus graft disease  
 PT in bone marrow and organ transplantation

XX Disclosure; Page 15; 47pp; English.

XX The present sequence represents a new peptide (PI) constituting a T-cell  
 CC epitope obtainable from the minor histocompatibility antigen HA-1. The  
 CC peptide is immunogenic and can be used as part of a vaccine. PI is used  
 CC as a medicine, to induce tolerance for transplants, prevent rejection  
 CC and/or graft versus host disease, or to treat (auto) immune diseases.  
 CC In particular it can be used with bone marrow transplantation, in the  
 CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency  
 CC diseases.

SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLKDDLEA 9  
 |||||  
 Db 1 VLKDDLEA 9

RESULT 3  
 ID AAW97375  
 AC AAW97375 standard; Protein; 9 AA.

XX AC AAW97375;

XX ET 13-MAY-1999 (first entry);

XX HA-1 H-allele sequence.

XX Intron: minor histocompatibility antigen HA-1; typing allele;  
 KW H allele; R allele; polymorphic nucleotide; HA-1 typing;  
 KW bone marrow transplant; severe aplastic anaemia; leukaemia;  
 KW immune deficiency disease; ss.

XX Homo sapiens.

XX WO9905313-A2.

XX 04-FEB-1999.

XX 23-JUL-1998; 98WO-EP04928.

XX 02-JUN-1998; 98EP-C870125.

XX 23-JUL-1997; 97EP-0202303.

XX (UYLE-); RIJKSUNIV LEIDEN.

XX Goulmy E;

XX WPI; 1999-142960/12.

XX Typing minor histocompatibility antigen HA-1 - by amplifying and  
 PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g.  
 PT detection of genetic aberrances

XX Claim 18; Fig 5; 59pp; English.

XX The present sequence represents part of the minor histocompatibility  
 CC antigen HA-1 H-allele. The specification describes methods for typing  
 CC alleles (preferably the H and R alleles) of the minor histocompatibility  
 CC antigen HA-1 in a sample, which comprise detecting polymorphic  
 CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The  
 CC methods can be used for HA-1 typing for bone marrow transplants, severe  
 CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as  
 CC detection of genetic aberrances. The probes and primers of the invention  
 CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be  
 CC used anti-idiotypic B cells and/or T cells and antibodies.

XX Sequence 9 AA;

Query Match 88.1%; Score 37; DB 20; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 9.3e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLKDDLEA 9  
 |||||  
 Db 1 VLKDDLEA 9

RESULT 4  
 AAW99196

```

ID AAW99196 standard; peptide; 9 AA.
XX
AC AAW99196;
XX
XX 20-MAY-1999 (first entry)
XX
DE Minor histocompatibility antigen HA-1 T-cell epitope #2.
XX
XX Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
KW diagnosis; aplastic anaemia; immune deficiency disease.
XX
XX Homo sapiens.
XX
XX WO9905174-A1.
XX
XX 04-FEB-1999.
XX
XX 23-JUL-1998; 98WO-NL00425.
XX
XX 23-JUL-1997; 97EP-0202303.
XX
XX (UYLE-) RIJKSUNIV LEIDEN.
XX
XX Engelhard VH, Goulmy EAJM, Hunt DF;
XX
XX WPI; 1999-153312/13.
XX
XX A new minor histocompatibility antigen, HA-1 - useful to treat
PT immune diseases and prevent rejection and host versus graft disease
PT in bone marrow and organ transplantation
XX
XX Claim 3; Page 32; 47pp; English.
XX
XX The present sequence represents a new peptide (P1) constituting a T-cell
CC epitope obtainable from the minor histocompatibility antigen HA-1. The
CC peptide is immunogenic and can be used as part of a vaccine. P1 is used
CC as a medicine, to induce tolerance for transplants, prevent rejection
CC and/or graft versus host disease, or to treat (auto) immune diseases.
CC In particular it can be used with bone marrow transplantation, in the
CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency
CC diseases.
XX
XX Sequence 9 AA;
XX
Query Match 88.1%; Score 37; DB 23; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLKDDLLLEA 9
Db 1 VLKDDLLLEA 9

RESULT 5
AAW97572
ID AAW97572 standard; peptide; 9 AA.
XX
XX AAW97572;
XX
XX 20-MAY-1999 (first entry)
XX
XX T-cell epitope from the minor histocompatibility antigen HA-1.
DE
XX T-cell epitope; minor histocompatibility antigen HA-1; vaccine;
KW transplant rejection; Graft-versus-Host Disease; autoimmune disease;
KW neoplastic haematopoietic cell.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Misc-difference 3 /note= "His or Arg"
FT

```

```

XX WO9905173-A1.
XX
XX 04-FEB-1999.
XX
XX 23-JUL-1998; 98WO-NL00424.
XX
XX 23-JUL-1997; 97EP-0202303.
XX
XX (UYLE-) RIJKSUNIV LEIDEN.
XX
XX Engelhard VH, Goulmy EAJM, Hunt DF;
XX
XX WPI; 1999-142855/12.
XX
XX Immunogenic peptide from minor histocompatibility antigen HA-1 -
PT useful for inducing tolerance to transplants and prevent rejection
PT or graft-versus-host disease
XX
XX Claim 1; Page 39; 57pp; English.
XX
XX The present sequence represents an immunogenic peptide constituting a
CC T-cell epitope, obtainable from the minor histocompatibility antigen
CC HA-1. The peptide can be used in vaccines or pharmaceutical formulations
CC as medicines to induce tolerance for transplants so as to prevent
CC rejection and/or Graft-versus-Host Disease, or to treat autoimmune
CC diseases. Neoplastic haematopoietic cells presenting the peptides, in
CC an HLA class I context, can be eliminated after specific recognition
CC of the peptides. The peptides can also be used to raise antibodies,
CC T-cell receptor, B- and T-cells.
XX
XX Sequence 9 AA;
XX
Query Match 95.7%; Score 36; DB 20; Length 9;
Best Local Similarity 88.9%; Pred. No. 9.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLKDDLLLEA 9
Db 1 VLKDDLLLEA 9

RESULT 6
AAW93195
ID AAW93195 standard; peptide; 9 AA.
XX
XX AAW93195;
XX
XX 20-MAY-1999 (first entry)
XX
XX Minor histocompatibility antigen HA-1 T-cell epitope #1.
DE
XX Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
KW diagnosis; aplastic anaemia; immune deficiency disease.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 3 /label= His, Arg
FT
XX WO9905174-A1.
XX
XX 04-FEB-1999.
XX
XX 23-JUL-1998; 98WO-NL00425.
XX
XX 23-JUL-1997; 97EP-0202303.
XX
XX (UYLE-) RIJKSUNIV LEIDEN.
XX
XX

```

Engelhard VH, Goulmy EAJM, Hunt DF;  
WPI: 1999-1533:2/13.  
A new minor histocompatibility antigen, HA-1 - useful to treat  
immune diseases and prevent rejection and host versus graft disease  
in bone marrow and organ transplantation  
PS  
PS Claim 1; Page 32; 47pp; English.  
XX  
XX The present sequence represents a new peptide (P1) constituting a T-cell  
CC epitope obtainable from the minor histocompatibility antigen HA-1. The  
CC peptide is immunogenic and can be used as part of a vaccine. P1 is used  
CC as a medicine, to induce tolerance for transplants, prevent rejection  
CC and/or graft versus host disease, or to treat (auto) immune diseases.  
CC In particular it can be used with bone marrow transplantation, in the  
CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency  
CC diseases.  
XX  
XX  
SQ Sequence 9 AA;  
Query Match 85.7%; Score 36; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 9.3e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VLRDLDLEA 9  
DB 1 VLKDDLEA 9  
RESULT 7  
AAW47427  
ID AAW47427 standard; peptide; 9 AA.  
XX  
XX  
AC AAW47427;  
XX  
XX  
DT 05-JUN-1998 (first entry)  
XX  
XX Prenyl diphosphate synthetase preserved region VI sequence.  
DE  
DE Prenyl diphosphate synthetase; polyprenyl diphosphate; vitamin K;  
KW ubiquinone; preserved region VI.  
XX  
XX Synthetic.  
OS  
XX EP812914-A2.  
PN  
XX 17-DEC-1997.  
PD  
XX  
XX 13-JUN-1997; 97EP-0:09692.  
PF  
XX 14-JUN-1996; 96JP-C:54441.  
PR  
XX (TCYT) TOYOTA JIDOSHA KK.  
PA  
XX Cho Y, Koike A, Koyama T, Muramatsu M, Ogura K;  
PI Shimizu N;  
PI  
XX WPI: 1998-034975/04.  
DR  
XX  
XX DNA encoding prenyl diphosphate synthetase subunit(s) - new  
PT Micrococcus prenyl diphosphate synthetase subunit polypeptide(s),  
PT and methods for preparing enzymes from subunit(s)  
XX  
XX Example 2; Page 26; 46pp; English.  
PS  
XX The present sequence from preserved region VI of prenyl diphosphate  
CC synthetase (PDS), was used in the preparation of primers for the  
CC amplification of PDS DNA.  
CC Substances synthesised by PDS, i.e. polyprenyl diphosphates, are  
CC precursors of physiologically active substances, e.g. vitamin K and  
CC ubiquinones.  
XX

SQ Sequence 9 AA;  
Query Match 64.1%; Score 27; DB 19; Length 9;  
Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LRDDLE 9  
DB 3 LRDDLE 9  
RESULT 8  
AAU71428  
ID AAU71428 standard; Peptide; 9 AA.  
XX  
XX AAU71428;  
AC  
XX  
XX 26-FEB-2002 (first entry)  
DT  
XX Human MHC molecule HLA-A\*11 binding 103P3E8 peptide #35.  
XX  
XX 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;  
KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;  
XX single chain monoclonal antibody; serum; blood; urine; tissue; human;  
XX chromosome 9q13-q21.  
OS  
XX Homo sapiens.  
XX  
XX WO200179557-A2.  
FN  
XX 25-OCT-2001.  
FD  
XX  
XX 12-APR-2001; 2001WO-JS12:51.  
FF  
XX  
XX 12-APR-2000; 2000US-196647P.  
PR  
XX (URCG-) UROGENESYS INC.  
PA  
XX  
XX Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DB;  
PI Cakobovits A;  
PI  
XX WPI: 2002-061976/08.  
DR  
XX  
XX Monitoring 103P3E8 gene products in sample from patient (suspected of)  
PT having cancer, useful for diagnosing, managing or treating cancers,  
PT e.g. prostate cancer, comprises determining presence of aberrant  
PT 103P3E8 gene products  
XX  
XX Disclosure; Page 90; 128pp; English.  
PS  
XX  
XX Sequences AAU71893-AAU71796 represent the 103P3E8-related protein and  
CC peptide fragments of the protein. 103P3E8 exhibits tissue specific  
CC expression in normal adult tissue, but it is also aberrantly expressed in  
CC many cancers including tumours of the prostate, bladder, kidney, colon,  
CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related  
CC protein and peptide fragments and specific PCR primers are therefore  
CC useful for diagnosing and treating cancer. A vector comprising a  
CC polynucleotide which encodes a single chain monoclonal antibody, that  
CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme  
CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,  
CC are both useful in the preparation of a composition for treating a  
CC patient with a cancer that expresses 103P3E8. The sequences can be used  
CC in diagnostic methods to monitor the level of 103P3E8 gene products in  
CC serum, blood, urine and tissue and to thereby detect the presence of  
CC cancerous cells.  
XX  
XX  
SQ Sequence 9 AA;  
Query Match 59.5%; Score 25; DB 23; Length 9;  
Best Local Similarity 62.5%; Pred. No. 9.3e+05;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2 LRDDLEA 9





```

XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response
PT in cancer patients
XX
XX
PS Claim 13; Page 311; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention.
XX
SQ Sequence 9 AA;
Query Match 54.8%; Score 23; DB 24; Length 9;
Best Local Similarity 71.4%; Pred.No. 9.3e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 LRDDLE 8
DB 2 LRDELE 8
RESULT 14
ABR19578
ID ABR19578 standard; Peptide; 9 AA.
XX
AC ABR19578;
XX
XX 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 184P3G10 HLA peptide #813.
XX
XX Human; cytostatic; vaccine; cancer; immune response; HLA;
XX human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 184P3G10 HLA peptide #813.
XX
XX Human; cytostatic; vaccine; cancer; immune response; HLA;
XX human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US11654.
XX
PR 10-APR-2001; 2001US-262739P.
XX
PR 10-APR-2001; 2001US-283112P.
XX
PR 25-APR-2001; 2001US-286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
XX Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;
XX Morrison K, Morrison RK, Raitano AB;
XX
WPI; 2003-075555/07.
XX
CC New composition comprising a substance that modulates the structure of
CC proteins and polynucleotides, useful for therapeutic, prognostic and
CC diagnostic reagents for eliciting cellular or humoral immune response
CC in cancer patients
XX
XX
PS Claim 13; Page 313; 1021pp; English.
XX

```

```

CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention.
XX
SQ Sequence 9 AA;
Query Match 54.8%; Score 23; DB 24; Length 9;
Best Local Similarity 71.4%; Pred.No. 9.3e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 LRDDLE 8
DB 2 LRDELE 8
RESULT 15
ABR19780
ID ABR19780 standard; Peptide; 9 AA.
XX
AC ABR19780;
XX
XX 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 184P3G10 HLA peptide #1015.
XX
XX Human; cytostatic; vaccine; cancer; immune response; HLA;
XX human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PT 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US11654.
XX
PR 10-APR-2001; 2001US-282739P.
XX
PR 10-APR-2001; 2001US-283112P.
XX
PR 25-APR-2001; 2001US-286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
XX Jakobovits A, Challita-Eid PM, Paris M, Ge W, Hubert RS;
XX Morrison K, Morrison RK, Raitano AB;
XX
WPI; 2003-075555/07.
XX
CC New composition comprising a substance that modulates the structure of
CC proteins and polynucleotides, useful for therapeutic, prognostic and
CC diagnostic reagents for eliciting cellular or humoral immune response
CC in cancer patients
XX
XX
PS Claim 13; Page 315; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present

```



CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
 CC from the invention.

XX  
 SQ Sequence 9 AA:

Query Match 54.8%; Score 23; DB 24; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 1; Mismatches 1; Indels 3; Gaps 0;

QY 2 LRDLLE 8  
 Do 2 LRGELE 8

Search completed: October 20, 2003, 13:45:48  
 Job time : 57 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:48:32 ; Search time 43 Seconds  
(without alignments)  
34,308 Million cell updates/sec

Title: SEQ1B  
Perfect score: 42  
Sequence: 1 vlrdillea 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 609560 seqs, 163917102 residues

Total number of hits satisfying chosen parameters: 60600

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubaa/PT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubaa/PT08\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubaa/JSC9A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	59.5	9	9	US-09-834-765-349
2	24	57.1	7	12	US-10-166-225A-170
3	23	54.8	9	9	US-09-834-765-132
4	21	50.0	7	9	US-09-989-789-208
5	21	50.0	7	9	US-09-989-789-211
6	21	50.0	7	9	US-09-989-789-216
7	21	50.0	7	11	US-09-990-186-208
8	21	50.0	7	11	US-09-990-186-211
9	21	50.0	7	11	US-09-990-186-216
10	21	50.0	7	11	US-09-989-994-208
11	21	50.0	7	11	US-09-989-994-211
12	21	50.0	7	11	US-09-989-994-216
13	21	50.0	9	12	US-09-833-203-54
14	20	47.6	5	12	US-10-166-225A-168
15	20	47.6	5	12	US-10-166-225A-169

16	20	47.6	5	12	US-10-166-225A-170	Sequence 170, App
17	20	47.6	5	12	US-10-166-225A-171	Sequence 171, App
18	20	47.6	5	12	US-10-166-225A-172	Sequence 172, App
19	20	47.6	5	12	US-10-166-225A-173	Sequence 173, App
20	20	47.6	7	12	US-10-271-708-8	Sequence 8, App
21	20	47.6	8	9	US-09-801-784-19	Sequence 19, App
22	20	47.6	8	9	US-09-801-784-20	Sequence 20, App
23	20	47.6	8	9	US-09-801-784-21	Sequence 21, App
24	20	47.6	8	9	US-09-801-784-22	Sequence 22, App
25	20	47.6	9	12	US-09-932-165-1239	Sequence 1239, App
26	20	47.6	9	12	US-09-935-384-104	Sequence 104, App
27	20	47.6	9	12	US-09-935-384-349	Sequence 349, App
28	20	47.6	9	12	US-09-935-384-515	Sequence 515, App
29	20	47.6	9	12	US-09-935-384-548	Sequence 548, App
30	20	47.6	9	12	US-09-935-384-620	Sequence 620, App
31	19	45.2	7	9	US-09-989-789-2954	Sequence 2954, App
32	19	45.2	7	9	US-09-989-789-3877	Sequence 3877, App
33	19	45.2	7	9	US-09-989-789-3914	Sequence 3914, App
34	19	45.2	7	9	US-09-989-789-3915	Sequence 3915, App
35	19	45.2	7	9	US-09-989-789-3916	Sequence 3916, App
36	19	45.2	7	9	US-09-989-789-3932	Sequence 3932, App
37	19	45.2	7	9	US-09-989-789-3941	Sequence 3941, App
38	19	45.2	7	11	US-09-990-186-2954	Sequence 2954, App
39	19	45.2	7	11	US-09-990-186-3877	Sequence 3877, App
40	19	45.2	7	11	US-09-990-186-3914	Sequence 3914, App
41	19	45.2	7	11	US-09-990-186-3915	Sequence 3915, App
42	19	45.2	7	11	US-09-990-186-3916	Sequence 3916, App
43	19	45.2	7	11	US-09-990-186-3932	Sequence 3932, App
44	19	45.2	7	11	US-09-990-186-3941	Sequence 3941, App
45	19	45.2	7	11	US-09-989-994-2954	Sequence 2954, App

#### ALIGNMENTS

RESULT 1  
US-09-834-765-349  
; Sequence 349, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Chailita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; PRIOR FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 349  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-349

Query Match 59.5%; Score 25; DB 9; Length 9;  
Best Local Similarity 62.5%; Pred No. 5.4e+05;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CY 2 LRDDLEA 9  
..|||  
DB 2 VKDLEA 9

RESULT 2  
US-10-271-708-16  
; Sequence 16, Application US/10271708

```

; Publication No. US20030162200A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: LIU, Xiaohai
; APPLICANT: BALASUBRAMANIAN, Shankar
; APPLICANT: PATES, Sachin D.
; APPLICANT: ISALAN, Mark
; TITLE OF INVENTION: METHODS FOR MODULATING TELOMERASE ACTIVITY
; FILE REFERENCE: SABI-036/0:US (8325-2014.01)
; CURRENT APPLICATION NUMBER: US/10/271,708
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 16
; LENGTH: 7
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: F2 residue
US-10-271-708-16

```

```

Query Match 57.1%; Score 24; DB 12; Length 7;
Best Local Similarity 83.3%; Pred. No. 5.4e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 RDDLE 8
DB 2 RDDLE 7

```

```

RESULT 3
US-09-834-765-132
; Sequence 132, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Mary Paris
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Ava Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; FILE REFERENCE: 129.6USU1
; CURRENT APPLICATION NUMBER: US/09/834,765
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-132

```

```

Query Match 54.8%; Score 23; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 DLLEA 9
DB 1 DLLEA 5

```

```

RESULT 4
US-09-989-789-208
; Sequence 208, Application US/0989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:

```

```

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: TRIPLETS BY ZINC FINGERS

```

```

; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 209
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example 2FP
US-09-989-789-208

```

```

Query Match 50.0%; Score 21; DB 9; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 RDDLE 8
DB 1 RDDLE 6

```

```

RESULT 5
US-09-989-789-211
; Sequence 211, Application US/0989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:

```

```

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: TRIPLETS BY ZINC FINGERS

```

```

; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 211
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example 2FP
US-09-989-789-211

```

```

Query Match 50.0%; Score 21; DB 9; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 RDDLE 8
DB 1 RDDLE 6

```

```

RESULT 6
US-09-989-789-216
; Sequence 216, Application US/0989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:

```

```

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: TRIPLETS BY ZINC FINGERS

```

```

; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 216
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example 2FP
US-09-989-789-216

```

```
Query Match          50.0%; Score 21; DB 9; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLLE 8
   |   |   |
Db 1 RSDLLQ 6

RESULT 7
US-09-990-186-208
; Sequence 208, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE OF INVENTION: TRIPLET BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 208
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-990-186-208

Query Match          50.0%; Score 21; DB 11; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLLE 8
   |   |   |
Db 1 RSDLLQ 6

RESULT 8
US-09-990-186-211
; Sequence 211, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE OF INVENTION: TRIPLET BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-990-186-211

Query Match          50.0%; Score 21; DB 11; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLLE 8
   |   |   |
Db 1 RSDLLQ 6

RESULT 9
US-09-990-186-216
; Sequence 216, Application US/09990186
; Publication No. US20030068675A1
```

```
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE OF INVENTION: TRIPLET BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 216
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-990-186-216

Query Match          50.0%; Score 21; DB 11; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLLE 8
   |   |   |
Db 1 RSDLLQ 6

RESULT 10
US-09-989-994-208
; Sequence 208, Application US/09989994
; Publication No. US20030104526A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE OF INVENTION: TRIPLET BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,994
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 208
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-994-208

Query Match          50.0%; Score 21; DB 11; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLLE 8
   |   |   |
Db 1 RSDLLQ 6

RESULT 11
US-09-989-994-211
; Sequence 211, Application US/09989994
; Publication No. US20030104526A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE OF INVENTION: TRIPLET BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,994
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 211
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: example ZFP  
US-09-989-994-211

Query Match 50.0%; Score 21; DB 11; Length 7;  
Best Local Similarity 66.7%; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLLE 8  
DB 1 RSDLLQ 6

RESULT 12  
US-09-989-994-216  
; Sequence 216, Application US/09989994  
; Publication No. US20030104526A  
; GENERAL INFORMATION:  
; APPLICANT: LUC, Quang  
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE  
; TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS  
; FILE REFERENCE: 8325-0011.20 / S11-US2  
; CURRENT APPLICATION NUMBER: US/09/989,994  
; CURRENT FILING DATE: 2001-11-20  
; NUMBER OF SEQ ID NOS: 4085  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 216  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP  
US-09-989-994-216

Query Match 50.0%; Score 21; DB 11; Length 7;  
Best Local Similarity 66.7%; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RDDLLE 8  
DB 1 RSDLLQ 6

RESULT 13  
US-09-833-203-54  
; Sequence 54, Application US/09833203  
; Publication No. US20030166277A1  
; GENERAL INFORMATION:  
; APPLICANT: Zauderer, Maurice  
; APPLICANT: Smith, Ernest S.  
; TITLE OF INVENTION: Targeted Vaccine Delivery Systems  
; FILE REFERENCE: 1821.002001  
; CURRENT APPLICATION NUMBER: US/09/833,203  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: US 60/196,472  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 54  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: C35 peptides  
US-09-833-203-54

Query Match 50.0%; Score 21; DB 12; Length 9;  
Best Local Similarity 80.0%; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLEA 9  
DB 1 RSDLLQ 6

QY 1 DLLEA 5

RESULT 14  
US-10-166-225A-168  
; Sequence 168, Application US/10-166225A  
; Publication No. US20030148416A1  
; GENERAL INFORMATION:  
; APPLICANT: BERRY, Alan  
; APPLICANT: BREITZEL, Werner  
; APPLICANT: HUMSELIN, Markus  
; APPLICANT: LOPEZ-CEBARRI, Rual  
; APPLICANT: MAYER, Anne F.  
; APPLICANT: YELISEEV, Alexei A.  
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION  
; FILE REFERENCE: C38435/121966  
; CURRENT APPLICATION NUMBER: US/10/166,225A  
; CURRENT FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 168  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Bradyrhizobium japonicum  
US-10-166-225A-168

Query Match 47.6%; Score 20; DB 12; Length 5;  
Best Local Similarity 60.0%; Pred. No. 5.4e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
DB 1 DDLLE 5

RESULT 15  
US-10-166-225A-169  
; Sequence 169, Application US/10166225A  
; Publication No. US20030148416A1  
; GENERAL INFORMATION:  
; APPLICANT: BERRY, Alan  
; APPLICANT: BREITZEL, Werner  
; APPLICANT: HUMSELIN, Markus  
; APPLICANT: LOPEZ-CEBARRI, Rual  
; APPLICANT: MAYER, Anne F.  
; APPLICANT: YELISEEV, Alexei A.  
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION  
; FILE REFERENCE: C38435/121966  
; CURRENT APPLICATION NUMBER: US/10/166,225A  
; CURRENT FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 169  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Rhizobium sp. strain NGR234  
US-10-166-225A-169

Query Match 47.6%; Score 20; DB 12; Length 5;  
Best Local Similarity 60.0%; Pred. No. 5.4e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
DB 1 DDLLE 5

Search completed: October 20, 2003, 14:01:22  
Job time: 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:43:56 ; Search time 20 Seconds  
(without alignments)  
19.04C Million cell updates/sec

Title: SEQ1B

Perfect score: 42

Sequence: 1 virddilea 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 77317

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.rep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B.COMB.rep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A.COMB.rep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B.COMB.rep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTUS.COMB.rep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.rep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY-ES

Result No.	Score	Query Match	Length	ID	Description
1	27	64.3	9	3	US-09-217-609A-11
2	27	64.3	9	3	US-08-873-235B-11
3	22	52.4	7	3	US-08-556-419-14
4	21	50.0	6	2	US-08-459-568-20
5	21	50.0	6	2	US-08-399-411-20
6	21	50.0	6	3	US-08-516-859A-20
7	21	50.0	6	4	US-09-586-472-20
8	21	50.0	6	4	US-09-528-706-20
9	21	50.0	8	2	US-08-459-568-75
10	21	50.0	8	2	US-08-399-411-75
11	21	50.0	8	2	US-08-539-432-1
12	21	50.0	8	3	US-08-516-859A-75
13	21	50.0	8	4	US-09-586-472-75
14	21	50.0	8	4	US-09-528-706-75
15	21	50.0	9	2	US-08-459-568-72
16	21	50.0	9	2	US-08-459-568-78
17	21	50.0	9	2	US-08-399-411-72
18	21	50.0	9	2	US-08-399-411-78
19	21	50.0	9	3	US-08-516-859A-72
20	21	50.0	9	3	US-08-516-859A-78
21	21	50.0	9	4	US-09-586-472-72
22	21	50.0	9	4	US-09-586-472-78
23	21	50.0	9	4	US-09-528-706-72
24	21	50.0	9	4	US-09-528-706-78
25	20	47.6	6	3	US-09-035-819-11
26	20	47.6	6	4	US-09-808-126-11
27	20	47.6	6	4	US-09-803-951-11

28 20 47.6 8 3 US-08-747-599A-47 Sequence 47, Appl  
29 20 47.6 8 3 US-08-747-599A-50 Sequence 50, Appl  
30 20 47.6 8 3 US-08-747-599A-51 Sequence 51, Appl  
31 20 47.6 9 4 US-08-634-332A-50 Sequence 50, Appl  
32 19 45.2 7 4 US-09-187-859-1186 Sequence 1186, Ap  
33 19 45.2 7 4 US-09-839-542B-1186 Sequence 1186, Ap  
34 19 45.2 8 4 US-09-187-859-1189 Sequence 1189, Ap  
35 19 45.2 8 4 US-09-183-266A-38 Sequence 38, Appl  
36 19 45.2 8 4 US-09-839-542B-1189 Sequence 1189, Ap  
37 19 45.2 9 1 US-08-467-081-20 Sequence 20, Appl  
38 19 45.2 9 1 US-08-414-417B-20 Sequence 20, Appl  
39 19 45.2 9 2 US-08-486-348A-20 Sequence 20, Appl  
40 19 45.2 9 2 US-08-468-545B-20 Sequence 20, Appl  
41 19 45.2 9 2 US-08-986-234-132 Sequence 102, App  
42 19 45.2 9 3 US-08-158-339A-1155 Sequence 1155, Ap  
43 19 45.2 9 3 US-08-466-680B-20 Sequence 20, Appl  
44 19 45.2 9 3 US-09-502-600-76 Sequence 76, Appl  
45 19 45.2 9 3 US-09-502-600-134 Sequence 134, App

ALIGNMENTS

RESULT 1  
US-09-217-609A-11  
Sequence 11, Application US/09217609A  
Patent No. 6071733  
GENERAL INFORMATION:  
APPLICANT: MURAMATSU, Masayoshi  
APPLICANT: KOIKE, Ayumi  
APPLICANT: OGURA, Kyoze  
APPLICANT: KOYAMA, Tanetoshi  
APPLICANT: SHIMIZU, Nacto  
APPLICANT: CHO, Yewin  
TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, NW - Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: US  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: WordPerfect 6.1 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/217,609A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/873,235  
FILING DATE: 11-Jun-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: COFFENETI, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 10235/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-1776  
TELEFAX: 202-429-0796  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-217-609A-11

Query Match 64.3%; Score 27; DB 3; Length 9;  
Best Local Similarity 57.1%; Pred. No. 2.5e+05;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

QY      2 LRDDLE 8
      3 IRDDILD 9

DB      2 LRDDLE 8
      3 IRDDILD 9

RESULT 2
US-08-873-235B-11
; Sequence 11, Application US/08973235B
; Patent No. 6174715
; GENERAL INFORMATION:
; APPLICANT: KYRUMATSU, Masayoshi
; APPLICANT: KOIKE, Ayumi
; APPLICANT: OGURA, Kyozo
; APPLICANT: KOYAMA, Tanetschi
; APPLICANT: SHIMIZU, Naoto
; APPLICANT: CHO, Yenwin
; TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, NW - Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
; SOFTWARE: WordPerfect 6.1 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,235B
; FILING DATE: 11-Jun-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 154441/1996
; FILING DATE: 14-Jun-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: TOFFENETTI, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 10235/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-1776
; TELEFAX: 202-429-0796
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-873-235B-11

Query Match      64.1%; Score 27; DB 3; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.5e+05;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 LRDDLE 8
      3 IRDDILD 9

DB      2 LRDDLE 8
      3 IRDDILD 9

RESULT 3
US-08-556-419-14
; Sequence 14, Application US/08556419C
; Patent No. 6093549
; GENERAL INFORMATION:
; APPLICANT: Ross, Christopher
; APPLICANT: Li, Xiao-Jiang
; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan
; APPLICANT: Lanahan, Anthony
; APPLICANT: Worley, Paul
; APPLICANT: Snyder, Solomon
; TITLE OF INVENTION: Huntingtin-associated protein
; FILE REFERENCE: 01107,5227;
; CURRENT APPLICATION NUMBER: US/08/556,419C
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Pse-Seq for Windows Version 3.0
; SEQ ID NO: 14
; LENGTH: 7
; TYPE: PPT
; ORGANISM: Rattus norvegicus
US-08-556-419-14

Query Match      52.4%; Score 22; DB 3; Length 7;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4 DDLE 8
      5 DDLE 5

DB      4 DDLE 8
      5 DDLE 5

RESULT 4
US-08-459-568-20
; Sequence 20, Application US/08459568
; Patent No. 5811304
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell, and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,568
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/399,411
; FILING DATE: 06-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-459-568-20

Query Match      50.3%; Score 21; DB 2; Length 6;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4 DDLE 8
      5 DDLE 5

DB      4 DDLE 8
      5 DDLE 5

```

```

RESULT 5
US-08-399-411-20
: Sequence 20, Application US/08399411
: Patent No. 5831008
: GENERAL INFORMATION:
: APPLICANT: Huang, Shi
: TITLE OF INVENTION: Retinoblastoma Protein - Interacting
: NUMBER OF SEQUENCES: 93
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Campbell & Flores LLP
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: F-ppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/399,411
: FILING DATE: 06-MAR-1995
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-LJ 1264
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-8949
: TELEFAX: (619) 535-9001
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
US-08-399-411-20

Query Match 50.0%; Score 21; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE S
Db 1 EDLE 5

RESULT 6
US-08-516-859A-20
: Sequence 20, Application US/08516859A
: Patent No. 6069231
: GENERAL INFORMATION:
: APPLICANT: Huang, Shi
: TITLE OF INVENTION: Retinoblastoma Protein - Interacting
: NUMBER OF SEQUENCES: 106
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Campbell & Flores LLP
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/516,859A
: FILING DATE: 18-AUG-1995
: CLASSIFICATION: 514

```

```

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/399,411
: FILING DATE: 06-MAR-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/292,683
: FILING DATE: 18-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-LJ 1776
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-8949
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
US-08-516-859A-20

Query Match 50.0%; Score 21; DB 3; Length 6;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8
Db 1 EDLE 5

RESULT 7
US-09-586-472-20
: Sequence 20, Application US/09586472
: Patent No. 6323335
: GENERAL INFORMATION:
: APPLICANT: Huang, Shi
: TITLE OF INVENTION: Retinoblastoma Protein - Interacting
: NUMBER OF SEQUENCES: 106
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Campbell & Flores LLP
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/586,472
: FILING DATE: 01-Jun-2000
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/528,706
: FILING DATE: 17-MAR-2000
: APPLICATION NUMBER: US 08/516,859
: FILING DATE: 18-AUG-1995
: APPLICATION NUMBER: US 08/399,411
: FILING DATE: 06-MAR-1995
: APPLICATION NUMBER: US 08/292,683
: FILING DATE: 18-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-LJ 4130
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-8949
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6 amino acids

```



```

;
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-586-472-20
Query Match          50.0%; Score 21; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0;

```

```

QY      4 DDLLE 8
DB      1 EDLLE 5

```

```

RESULT 8
US-09-528-706-20
; Sequence 20, Application US/09528706
; Patent No. 6468985
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/528,706
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/516,859
; FILING DATE:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,683
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-528-706-20

```

```

Query Match          50.0%; Score 21; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 DDLLE 8
DB      1 EDLLE 5

```

```

RESULT 9
US-08-459-568-75
; Sequence 75, Application US/08459568
; Patent No. 5811304
; GENERAL INFORMATION:

```

```

; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,568
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-459-568-75

```

```

Query Match          50.0%; Score 21; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 DDLLE 8
DB      1 EDLLE 5

```

```

RESULT 10
US-08-399-411-75
; Sequence 75, Application US/08399411
; Patent No. 5811308
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,411
; FILING DATE: 06-MAR-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.

```

```

; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-399-411-75

```

```

Query Match          50.0%; Score 21; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 DDLLE 8
DB 1 EDLLE 5

```

```

RESULT 11
US-08-539-432-1
; Sequence 1, Application US/08519432
; Patent No. 5872210
; GENERAL INFORMATION:
; APPLICANT: MEDABALIM, JOHN L.
; TITLE OF INVENTION: TRANSFRAME INHIBITORY
; TITLE OF INVENTION: ELEMENT OF VIRAL
; TITLE OF INVENTION: PROTEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,432
; FILING DATE: 05-OCT-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4194
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4900
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-539-432-1

```

```

Query Match          50.0%; Score 21; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 LRDDL 6
DB 2 LRDDL 6

```

```

RESULT 12
US-08-516-859A-75

```

```

; Sequence 75, Application US/09516659A
; Patent No. 6669231
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/516,859A
; FILING DATE: 18-AUG-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,683
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-516-859A-75

```

```

Query Match          50.0%; Score 21; DB 3; Length 8;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 DDLLE 8
DB 1 EDLLE 5

```

```

RESULT 13
US-08-586-472-75
; Sequence 75, Application US/09586472
; Patent No. 6323335
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/586,472  
FILING DATE: 01-Jun-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/528,706  
FILING DATE: 17-MAR-2000  
APPLICATION NUMBER: US 08/516,859  
FILING DATE: 18-AUG-1994  
APPLICATION NUMBER: US 08/399,411  
FILING DATE: 06-MAR-1995  
APPLICATION NUMBER: US 08/292,683  
FILING DATE: 18-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 4130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-8949  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 75:  
US-09-586-472-75

Query Match 50.0%; Score 21; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0;

QY 4 DDLLE 8  
DB 1 EDLLE 5

RESULT 14  
US-09-528-706-75  
Sequence 75, Application US/09528706  
Patent No. 6468985  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/528,706  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/516,859  
FILING DATE:  
APPLICATION NUMBER: US 08/292,683  
FILING DATE: 18-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1776  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-528-706-75

Query Match 50.0%; Score 21; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0;

QY 4 DDLLE 8  
DB 1 EDLLE 5

RESULT 15  
US-08-459-568-72  
Sequence 72, Application US/08459568  
Patent No. 5811304  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,568  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/399,411  
FILING DATE: 06-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 72:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-459-568-72

Query Match 50.0%; Score 21; DB 2; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0;

QY 4 DDLLE 8  
DB 1 EDLLE 5

Search completed: October 20, 2003, 13:50:07  
Job time: 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: October 20, 2003, 13:11:00 ; Search time 38 Seconds  
(without alignments)  
22.777 Million cell updates/sec

Title: US-09-489-760A-1  
Perfect score: 38  
Sequence: 1 VLXDDLLLEA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 789

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR 76:  
1: P121:  
2: P122:  
3: P123:  
4: P124:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	42.1	7	S68304	kucolin, 75k chain
2	16	42.1	8	PC4131	hypothetical prote
3	14	36.9	7	S20446	elastase - Pseudom
4	13	34.2	8	A61328	trypsin (EC 3.4.21
5	13	34.2	9	A60427	macrophage cytotox
6	13	34.2	9	PH0342	T-cell receptor be
7	12	31.6	4	I40897	biotin A - Citroba
8	12	31.6	5	F70621	T-cell receptor be
9	12	31.6	5	F70679	T-cell receptor be
10	12	31.6	6	B35640	cerebellar degener
11	12	31.6	6	PT0333	T-cell receptor be
12	12	31.6	7	A34026	acetylcholinestera
13	12	31.6	7	B39040	caldesmonin, fas
14	12	31.6	7	PT0628	T-cell receptor be
15	12	31.6	7	PT0722	T-cell receptor be
16	12	31.6	7	PT0576	T-cell receptor be
17	12	31.6	8	PT0557	T-cell receptor be
18	12	31.6	9	PS0353	glycine cleavage s
19	12	31.6	9	PH0108	late Gl-69 protein
20	12	31.6	9	PT0562	T-cell receptor be
21	12	31.6	9	B30572	T-cell receptor be
22	11	28.9	6	T11779	phosphoglycerate t
23	11	28.9	7	S25266	biLE protein - Esc
24	11	28.9	8	S22428	chitin-binding pro
25	11	28.9	8	B33039	158K exoantigen -
26	11	28.9	8	S69165	ferredoxin a2 - Ja
27	11	28.9	9	FW0002	chlorophyll a/b-bi
28	11	28.9	9	S66419	tetrameric protein
29	11	28.9	9	A12872	transaldolase (EC

30	11	28.9	9	2	A61386	macrophage inhibit
31	10	26.3	4	2	A48360	gamma subunit of P
32	10	26.3	4	2	A26299	protein-glutamine
33	10	26.3	6	2	H48394	glycoprotein compo
34	10	26.3	6	2	I48126	alpha-tubulin - Ch
35	10	26.3	7	2	S78024	ribosomal protein
36	10	26.3	7	2	A12016	formylglycinamide
37	10	26.3	7	2	C56793	platelet glycoprot
38	10	26.3	7	2	S29735	polyphosphate-gluc
39	10	26.3	7	2	PC2370	probable H+-transp
40	10	26.3	7	2	T09512	NADH2 dehydrogenas
41	10	26.3	8	2	A32523	peptidyl-di-peptida
42	10	26.3	8	2	T10077	hypothetical prote
43	10	26.3	8	2	PT0298	Ig heavy chain CAD
44	10	26.3	8	2	A42057	fibroblast growth
45	10	26.3	9	2	A35768	T-cell receptor al

ALIGNMENTS

RESULT 1  
S68304  
kucolin, 75k chain - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S68004  
R:EDGAR, P.F.  
PES Lett. 375, 159-161, 1995  
A>Title: Kucolin, a new corticosteroid-binding protein from human plasma with structur  
A:Reference number: S68004; XUID:96087107; PMID:7498469  
A:Accession: S68304  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <EDG>

Query Match 42.1%; Score 16; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred.No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6  
DB 4 DDL 6

RESULT 2  
PC4131  
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)  
C:Species: Pseudomonas aeruginosa  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: PC4131  
R:Kawasaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.  
Gene 167, 87-91, 1995  
A>Title: Sequencing and characterization of the downstream region of the genes encodin  
y for biosynthesis of heme d1.  
A:Reference number: JC4552; XUID:96144254; PMID:8566817  
A:Accession: PC4131  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-8 <KAW>  
A:Cross-references: DBJ:D50473; NID:g1217594  
A>Note: this ORF is not annotated in GenBank entry PSENIRC, release 113.0

Query Match 42.1%; Score 16; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred.No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDL 6  
DB 2 DDL 4

RESULT 3

```

S23446
elastase - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 24-Jul-1997
C:Accession: S20446
R/Kessler, E.; Saffrin, M.; Peretz, M.; Burstein, Y.
FEBS Lett. 299, 291-293, 1992
A:Title: Identification of cleavage sites involved in proteolytic processing of Pseudomonas elastase
A:Reference number: S20446; MUID:92183956; PMID:1544509
A:Accession: S20446
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <KES>

Query Match 36.9%; Score 14; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLE 8
DB 3 DLD 6

RESULT 4
A61328
trypsin (EC 3.4.21.4) precursor - minke whale (fragment)
C:Species: Balaenoptera acutorostrata (minke whale, lesser rorqual)
C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
C:Accession: A61328
R/Bricteux-Gregoire, S.; Schyns, R.; Florin, M.; Emmens, M.; Welling, G.M.; Beintema, J.
Biochim. Biophys. Acta 366, 244-255, 1975
A:Title: N-terminal amino acid sequence of trypsinogen from the lesser rorqual, Balaenoptera acutorostrata
A:Reference number: A61328; MUID:75146765; PMID:1125273
A:Accession: A61328
A:Molecule type: protein
A:Residues: 1-8 <BRI>
C:Keywords: hydrolase; protein digestion; serine proteinase; zymogen
F:1-8/Domain: activation peptide #status experimental <APT>

Query Match 34.2%; Score 13; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXD 5
DB 3 LDD 6

RESULT 5
A60427
macrophage cytotoxicity-inducing factor, 29K - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 18-Jun-1993
C:Accession: A60427
R/Jones, C.M.; Prince, C.A.; Williams, J.S.
Exp. Hematol. 19, 704-709, 1991
A:Title: Purification and amino acid analysis of a human macrophage cytotoxicity-inducing factor
A:Reference number: A60427; MUID:91372335; PMID:1929970
A:Accession: A60427
A:Molecule type: protein
A:Residues: 1-9 <JON>
A:Note: the sequence from the text on page 706 is inconsistent with that from page 708
C:Keywords: cytokine

Query Match 34.2%; Score 13; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLD 4
DB 4 VLD 7

RESULT 6
P30942
T-cell receptor beta chain V-D-J region (clone 13) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C:Accession: P30942
R/Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
C. Exp. Med. 174, 1467-1476, 1991
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic arthritis
A:Reference number: P30942; MUID:92078857; PMID:1836012
A:Accession: P30942
A:Molecule type: mRNA
A:Residues: 1-9 <COL>
A:Experimental source: complete Freund's adjuvant-immunized lymph node
A:Note: the authors translated the codon TGC for residue 2 as Ala
C:Keywords: T-cell receptor

Query Match 34.2%; Score 13; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLE 8
DB 5 LLE 7

RESULT 7
I40697
Biotin A - Citrobacter freundii (fragment)
C:Species: Citrobacter freundii
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-1996
C:Accession: I40697
R/Shaban, D.; Campbell, A.
Gene 67, 203-211, 1988
A:Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter freundii
A:Reference number: I40697; MUID:8906280; PMID:2971595
A:Accession: I40697
A:Status: preliminary; translated from GS/EXBL/DCBJ
A:Molecule type: DNA
A:Residues: 1-4 <RES>
A:Cross-references: GB:M21922; NID:G144434

Query Match 31.6%; Score 12; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5
DB 3 DD 4

RESULT 8
P30601
T-cell receptor beta chain V-D-J region (120-2K) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: P30601; P30617; P30694
R/Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Functional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: P30601; MUID:91277601; PMID:1711558
A:Accession: P30601
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FEE>
A:Experimental source: newborn thymus, strain BALB/c, clone 120-2K
A:Accession: P30617
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FEJ>
A:Experimental source: newborn thymus, strain BALB/c, 120-2CA
A:Accession: P30694

```

```

A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FEZ>
A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-1H
C:Keywords: T-cell receptor

Query Match      31.6%; Score 12; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 DD 5
      ||
Db      4 DD 5

RESULT 9
PT0679
T-cell receptor beta chain V-D-J region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0679; PT0708
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991.
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0533
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEZ>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match      31.6%; Score 12; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 DD 5
      ||
Db      4 DD 5

RESULT 12
A34026
acetylcholinesterase (EC 3.1.1.7); 5.6S form - Pacific electric ray (fragment)
C:Species: Torpedo californica (Pacific electric ray)
C:Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 08-Nov-1996
C:Accession: A34026
R:Gibney, G.; MacPhee-Quigley, K.; Thompson, B.; Vedvick, T.; Low, M.G.; Taylor, S.S.;
J. Biol. Chem. 263, 1140-1145, 1988
A:Title: Divergence in primary structure between the molecular forms of acetylcholines
A:Reference number: A34026; MUID:8808723; PMID:3335534
A:Accession: A34026
A:Molecule type: protein
A:Residues: 1-7 <GIB>
C:Keywords: alternative splicing; carboxylic ester hydrolase

Query Match      31.6%; Score 12; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      6 LLEA 9
      |||
Db      1 LDNA 4

RESULT 13
B39040
calsequestrin, fast skeletal muscle - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 23-Feb-1997
C:Accession: B39040
R:Gala, S.E.; Jones, L.R.
J. Biol. Chem. 266, 391-398, 1991
A:Title: Phosphorylation of cardiac and skeletal muscle calsequestrin isoforms by case
A:Reference number: A39040; MUID:9109153; PMID:1985907
A:Accession: B39040
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <CAL>
C:Keywords: phosphoprotein; skeletal muscle

Query Match      31.6%; Score 12; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 DD 5
      |||
Db      1 DD 2

RESULT 14
PT0533
T-cell receptor beta chain V-D-J region (126-1AA) - mouse (fragment)

```

PT0628  
T-cell receptor beta chain V-D-J region (111-1AG) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0628  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0528  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-7 <FEE>  
A:Experimental source: newborn thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 31.6%; Score 12; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DD 5  
||  
Db 4 DD 5

RESULT 15  
PT0722  
T-cell receptor beta chain V-D-J region (135-1G) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0722  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0722  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-7 <FEE>  
A:Experimental source: newborn thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 31.6%; Score 12; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DD 5  
||  
Db 4 DD 5

Search completed: October 20, 2003, 13:16:33  
Job time : 40 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:03:44 ; Search time 24 Seconds  
(without alignments)  
17.635 Million cell updates/sec

Title: US-09-489-760A-1  
Perfect score: 38  
Sequence: VLXDDLEA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026755 residues

Total number of hits satisfying chosen parameters: 251

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Swissprot\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	31.6	5	1 B10A CITR	P13371 citrobacter
2	12	31.6	9	1 NSK1 SARBU	P41492 sarcophaga
3	12	31.6	9	1 R342 LITRU	P62075 litorea rub
4	12	31.6	9	1 UH42 HUMAN	P40329 homo sapien
5	11	28.9	6	1 TR22 PSEPU	P36414 pseudomonas
6	11	28.9	9	1 PTSP BOMMO	P82033 bombyx mori
7	11	28.9	9	1 TAL1 PICJA	P17440 picchia jadi
8	11	28.9	9	1 UF02 MOUSE	P38640 mus musculus
9	10	26.3	7	1 PP42 LYCES	P83379 lycopersico
10	10	26.3	9	1 LMIP LOCMI	P31799 locusta mig
11	10	26.3	9	1 UL4H HUMAN	P31934 homo sapien
12	9	23.7	7	1 AL12 CARMA	P81805 carcinus ma
13	9	23.7	7	1 FARI HELTI	P41871 heliosoma tr
14	9	23.7	7	1 FARB CALVO	P41866 calliphora
15	9	23.7	7	1 UF03 MOUSE	P38641 mus musculus
16	9	23.7	8	1 FARB CALVO	P41863 calliphora
17	9	23.7	8	1 WPI FERAT	P83395 perkinsus a
18	9	23.7	9	1 FIBB ERYPA	P19346 erythrocebu
19	9	23.7	9	1 IPYR RHOVI	P82992 rhodopsin
20	9	23.7	9	1 PPH1 LYCES	P83380 lycopersico
21	9	23.7	9	1 UL4E HUMAN	P31931 homo sapien
22	8	21.1	5	1 TRM3 ECOLI	P13973 escherichia
23	8	21.1	6	1 ACPH RABIT	P25154 oryctolagus
24	8	21.1	7	1 GFRE MOUSE	P99025 mus musculus
25	8	21.1	8	1 ACT CARMA	P80709 carcinus ma
26	8	21.1	8	1 AL12 CARMA	P81815 carcinus ma
27	8	21.1	8	1 AL18 CARMA	P81821 carcinus ma
28	8	21.1	8	1 CAD1 ENTPA	P13268 enterococcu
29	8	21.1	8	1 COW2 CONCP	P58785 conus purpu
30	8	21.1	8	1 PLP BRANA	P81707 brassica ra
31	8	21.1	9	1 BS43 SERPL	P83375 serratia pl
32	8	21.1	9	1 BUK CLOPA	P81337 clostridium
33	8	21.1	9	1 FARI CALVO	P41856 calliphora

34	8	21.1	9	1 FARB CALVO	P41857 calliphora
35	8	21.1	9	1 FARB CALVO	P41858 calliphora
36	8	21.1	9	1 FARB CALVO	P41859 calliphora
37	8	21.1	9	1 FARB CALVO	P41860 calliphora
38	8	21.1	9	1 FARB CALVO	P41861 calliphora
39	8	21.1	9	1 FARB CALVO	P41862 calliphora
40	8	21.1	9	1 FARB CALVO	P41865 calliphora
41	8	21.1	9	1 FRF1 SARBU	P83350 sarcophaga
42	8	21.1	9	1 LFCA STAAU	P36894 staphylococ
43	8	21.1	9	1 TAL3 PICJA	P17441 picchia jadi
44	7	18.4	3	1 LUXE VIBFI	P24272 vibrio fisc
45	7	18.4	5	1 RE21 LITRU	P82071 litorea rub

## ALIGNMENTS

RESULT 1				
ID	B10A_CITR	STANDARD;	PRT;	5 AA.
AC	P13371			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase			
DE	(EC 2.6.1.62) (7,8-diamino-9-oxononanoic acid aminotransferase) (DAPA			
DE	aminotransferase) (fragment).			
GN	B10A			
OS	Citrobacter freundii			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Citrobacter			
OX	NCBI_TaxID=546;			
RA	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89006280; PubMed=2971595;			
RA	Shuan E., Campbell A;			
RT	"Transcriptional regulation and gene arrangement of Escherichia coli,			
RT	Citrobacter freundii and Salmonella typhimurium biotin operons";			
RL	Gene 67:203-211(1986).			
CC	-!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-			
CC	oxononanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-			
CC	diaminononanoate.			
CC	-!- COFACTOR: Pyridoxal phosphate.			
CC	-!- PATHWAY: Biotin biosynthesis.			
CC	-!- SUBUNIT: Homodimer.			
CC	-!- SIMILARITY: Belongs to class-III of pyridoxal-phosphate-dependent			
CC	aminotransferases.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
EMBL; M21922; -; NOT_ANNOTATED_CDS.				
PIR; I40697; I40697.				
DR InterPro; IPR005814; Aminotrans_3.				
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; PARTIAL.				
KW Biotin biosynthesis; Transferase; Aminotransferase;				
KW Pyridoxal phosphate.				
FT NON_TER 5				
SQ SEQUENCE 5 AA; 582 MW; 6AAAB1B1A6F00000 CRC64;				

Query Match		31.6%	Score 12;	DB 1;	Length 5;
Best Loca. Similarity		100.0%	Pred. No. 1.3e+05;		
Matches		2;	Conservative	0;	Mismatches 0; Gaps 0;
QY		4 DD 5			
Db		4 DD 5			



```

CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC -!- PTM: Rubellidin 4.2 seems to differ from Rubellidin 4.3 by its C-
CC terminal amidation.
CC -!- MASS SPECTROMETRY: MW=883; METHOD=FAB.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 5 9 AMIDATION.
SQ SEQUENCE 9 AA: 594 MW: 202577205AA72728 CRC64;

Query Match 31.6%; Score 12; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 3;

QY 5 DIL 7
DB 5 DIL 7

RESULT 4
JHA2_HUMAN
ID JHA2_HUMAN STANDARD; PRT; 9 AA.
AC P40929;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein from 2D-page of heart (Spot 5603) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett C.N., Wheeler C.H., Baker C.S., Yaccub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994."
RL Electrophoresis 15:1459-1465(1994).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.0, ITS MW IS: 55.3 kDa.
FT NON_TER 9
SQ SEQUENCE 9 AA: 1104 MW: 88745185B5901B2CA CRC64;

Query Match 31.6%; Score 12; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DD 5
DB 8 DD 9

RESULT 5
TRP_PSEPU
ID TRP_PSEPU STANDARD; PRT; 6 AA.
AC P36414;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE TrpA operon transcriptional activator (Fragment).
GN TRP.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC S-RAIN=PpG1 C1S;
RX MEDLINE=89335826; PubMed=2533057;
RA Ebexly L., Crawford I.P.;
RT "DNA sequence of the tryptophan synthase genes of Pseudomonas
RT putida."
RL Biochimie 71:521-531(1989).
CC -!- FUNCTION: ACTIVATES THE EXPRESSION OF THE TRPBA GENES ENCODING

```

CC THE TWO TRYPTOPHAN SYNTHASE SUBUNITS. IN THE ABSENCE OF THE  
CC INDUCER (INDOLEGLYCEROL PHOSPHATE), TRYPI BINDS UPSTREAM OF THE  
CC TRPAB OPERON, OVERLAPPING ITS OWN PROMOTER REGION.  
CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EVBL: X13299; CAA31660.1; --  
CC InterPro: IPR000847; HTH\_LYSR.  
CC PROSITE: PS00044; HTH\_LYSR\_FAMILY; PARTIAL.  
CC Tryptophan biosynthesis; Transcription regulation; Activator;  
CC DNA-binding. 6 6  
CC NON\_TER 6 6  
CC SEQUENCE 6 AA; 683 MW; 77672AA:EDD6F000 CRC64;

Query Match 28.9%; Score 11; DB 1; Length 6;  
Best Local Similarity 40.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LXDDY 6  
Db 1 MAHDL 5

RESULT 6  
PSP\_BOMMO STANDARD; PRT; 9 AA.  
AC P8203;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Prothoracicotropic peptide (Bom-PSP).  
OS Bombyx mori (Silk moth).  
CC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
CC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
CC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=C145 X N140; T-SSJUE=Brain;  
RX MEDLINE=20002634; PubMed=10531308;  
RA Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara M., Nagata S.,  
RA Katoaka H.,  
RT "Identification of a prothoracicotropic peptide in the larval brain of  
RT the silkworm, Bombyx mori.";  
RL J. Biol. Chem. 274:31169-31173 (1999).  
RN [2]  
RP ERRATUM.  
RA Hua Y.-J., Tanaka Y., Nakamura K., Sakakibara M., Nagata S.,  
RA Katoaka H.,  
RL J. Biol. Chem. 275:9892-9892 (2000).  
CC -1- FUNCTION: Inhibits ecdysteroid biosynthesis in the prothoracic  
CC gland.

CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- DEVELOPMENTAL STAGE: EARLY FIFTH INSTAR.  
KW Hormone; Amidation.  
FT MOD\_RES 9 9 AMIDATION.  
SQ SEQUENCE 9 AA; 1090 MW; 3878C5B4472AB6C3 CRC64;

Query Match 28.9%; Score 11; DB 1; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 DLLEA 9  
Db 4 DLMSA 8

RESULT 7  
TALL\_PICUA STANDARD; PRT; 9 AA.  
AC P17430;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transaldolase I (EC 2.2.1.2) (Fragment).  
OS Pichia jadinii (Yeast) (Candida utilis).  
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
CC Saccharomycetales; Saccharomycetaceae; Pichia.  
OX NCBI\_TaxID=4903;  
RN [1]  
RP SEQUENCE.

RX MEDLINE=77110646; PubMed=556924;  
RA Sun S.C., Joris L., Tsolas O.;  
RT "Purification of crystallization of transaldolase isozyme I and  
RT evidence for different genetic origin of isozymes I and III in  
RT Candida utilis.";  
RL Arch. Biochem. Biophys. 178:69-78 (1977).

CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF  
CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.  
CC -1- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde  
CC 3-phosphate = D-erythrose 4-phosphate + D-fructose 6-phosphate.  
CC -1- PATHWAY: Pentose phosphate pathway; nonoxidative part.  
CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY. SUBFAMILY 1.  
CC PIR: A12872; A12872.  
CC InterPro: IPR001585; Transaldolase.  
CC PROSITE: PS00959; TRANSALDOLASE\_2; PARTIAL.  
CC PROSITE: PS1054; TRANSALDOLASE\_1; PARTIAL.  
KW Transferase; Pentose shunt.  
FT NON\_TER 1 1  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 1008 MW; 274P31AF0E5:E058 CRC64;

Query Match 28.9%; Score 11; DB 1; Length 9;  
Best Local Similarity 50.3%; Pred. No. 1.3e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 DLLL 7  
Db 6 DLLL 9

RESULT 8  
UF02\_XOUSE STANDARD; PRT; 9 AA.  
AC P36640;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Unknown protein from 2D-page of fibroblasts (P32) (Fragment).  
OS Mus musculus (Mouse).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.

RX TISSUE=Fibroblast;  
RX MEDLINE=95009907; PubMed=7523108;  
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;  
RT "Separation and sequencing of familial and novel murine proteins  
RT using preparative two-dimensional gel electrophoresis.";  
RS Electrophoresis 15:735-745 (1994).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 6.0. ITS MW IS: 32 kDa.  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 1102 MW; 7E73EAB6D05B1AAB CRC64;

Query Match 28.9%; Score 11; DB 1; Length 9;  
Best Local Similarity 20.0%; Pred. No. 1.3e+05;

Matches 1; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXXDL 6  
DB 1 MEJEI 5

## RESULT 9

PPH2\_LYCES STANDARD; PRT; 7 AA.  
AC P83379;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Purple acid phosphatase isozyme LesAP2 (EC 3.1.3.2) (Fragment).  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; Lamiales; Solanales; Solanaceae; Solanum.  
NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE, CATALYTIC ACTIVITY, SUBUNIT, SUBCELLULAR LOCATION, AND  
GP GLYCOSYLATION.  
STRAIN=cv. MoneyMaker; TISSUE=Seed;  
RX MEDLINE=22361242; PubMed=12473124;  
RA Bozzo G.G., Raghothama K.G., Flaxton W.C.;  
RT "Purification and characterization of two secreted purple acid  
phosphatase isozymes from phosphate-starved tomato (Lycopersicon  
esculentum) cell cultures.";  
RL Eur. J. Biochem. 269:6278-6286(2002).  
CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = a:  
CC alcohol + phosphate.  
CC -1- SUBUNIT: Monomer.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- PTM: Glycosylated.  
CC -1- MISCELLANEOUS: In L. esculentum there are at least two isozymes of  
purple acid phosphatase.  
KW Hydroxylase; Glycoprotein.  
FT NON\_TER 1 1  
FT NON\_TER 7 7  
SQ SEQUENCE 7 AA; 810 MW; 672AA862C9C729AC CRC64;

Query Match 26.3%; Score 10; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DL 6  
DB 6 DL 7

## RESULT 10

LMIP\_LGCM STANDARD; PRT; 9 AA.  
AC P31799;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-OCT-1993 (Rel. 27, Last annotation update)  
DE Locustaminohibiting peptide (LOM-MIP).  
OS Locusta migratoria (Migratory locust).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
NCBI\_TaxID=7004;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92179466; PubMed=1796179;  
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;  
RT "Isolation, identification and synthesis of locustaminohibiting  
peptide (LOM-MIP), a novel biologically active neuropeptide from  
Locusta migratoria.";  
RL Regul. Pept. 36:111-119(1991).

-1- FUNCTION: SUPPRESSES SPONTANEOUS CONTRACTIONS OF THE HINDGUT AND

CC -1- TISSUE SPECIFICITY: NEURONS LOCATED IN TWO VENTRAL CELL CLUSTERS  
IN THE SUBSCERPHAL GANGLION.  
CC PIR; A60065; AKQIM.  
KW Amidation; Neuropeptide.  
FT MOD\_RES 9  
SQ SEQUENCE 9 AA; 1060 MW; 387D7DD4472ABEC3 CRC64;

## AMIDATION

Query Match 26.3%; Score 10; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DL 6  
DB 4 DL 5

## RESULT 11

ULAH\_HUMAN STANDARD; PRT; 9 AA.  
ID ULAH\_HUMAN  
AC P31934;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 26-OCT-2001 (Rel. 40, Last annotation update)  
DE Unknown protein from 2D-page of liver tissue (Spot 94) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX TISSUE=Liver;  
RX MEDLINE=94147969; PubMed=8313870;  
RA Hughes G.J., Prutiger S., Paquet N., Pasquall C., Sanchez J.-C.,  
RA Tissot J.-D., Bairoch A., Appel R.D., Hochstrasser D.F.;  
RT "Human liver protein map: update 1993.";  
RL Electrophoresis 14:1216-1222(1993).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
PROTEIN IS: 5.5, IFS MW IS: 12 KDa.  
CC SWISS-2DPAGE; P31934; HUMAN.  
FT NON\_TER 1 1  
FT NON\_TER 3 3  
FT VARIANT 1 3  
FT VARIANT 3 3  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 990 MW; F1CC7B172AB6DDCA CRC64;

Query Match 26.3%; Score 10; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DL 6  
DB 5 DL 6

## RESULT 12

ALL2\_CARNA STANDARD; PRT; 7 AA.  
ID ALL2\_CARNA  
AC P81805;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Carcinustatin 2.  
OC Carcinus naenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocythera; Brachyura;  
OC Subrachyura; Portunodea; Portunidae; Carcinus.  
NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE.  
RX TISSUE=Cerebral ganglion, and Thoracic ganglion;  
RX MEDLINE=98121193; PubMed=9461295;  
RA Dave H., Connser A.H., Maestri J.-L., Scott A.G., Jares P.P.,

```

RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RL allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 77C MW; 672879CDBSD570 CRC64;

Query Match 23.7%; Score 9; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EA 9
DB 1 EA 2

RESULT 13
FARI_HELTI STANDARD; PRT; 7 AA.
AC P41871;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRamide-like neuropeptide GDFPLRF-amide.
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Planorbidae; Helisoma.
OX NCBI_TaxID=27815;
RN [-];
RP SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRamide-related peptides from the kidney of the snail, Helisoma
RL trivolvis.";
RL Peptides 15:31-36(1994).
CC -1- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING
CC THE KIDNEY, MANTLE AND SKIN.
CC -1- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA310 CRC64;

Query Match 23.7%; Score 9; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 DDLL 7
DB 2 DPFL 5

RESULT 14
FARB_CALVO STANDARD; PRT; 7 AA.
AC P41866;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliFMRamide 11.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Cecidoidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1];
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;

```

```

RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld C.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliFMRamides) from the blowfly
RT Calliphora vomitoria";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
CC FAMILY.
DR PIR: B44787; B44787.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 926 MW; 69D40699C44AB700 CRC64;

Query Match 23.7%; Score 9; DB 1; Length 7;
Best Local Similarity 25.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 DDLL 7
DB 2 DNFM 5

RESULT 15
UF03_MOUSE STANDARD; PRT; 7 AA.
AC P38641;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Unknown protein from 2D-page of fibroblasts (P36; (Fragment)).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE.
RC TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familiar and novel murine proteins
RT using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -1- M-SC3LJANE05: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.1, ITS MW IS: 36 kDa.
FT NON_TER 7 7
SQ SEQUENCE 7 AA; 842 MW; 6AA72b1DDB1B1180 CRC64;

Query Match 23.7%; Score 9; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EA 9
DB 3 EA 4

Search completed: October 20, 2003, 13:14:07
Job time : 26 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:10:35 ; Search time 93 Seconds  
(without alignments:  
24.973 Million cell updates/sec

Title: US-09-489-760A-1  
Sequence: 1 VLXDDLLEA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 775

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mbc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriopl:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	16	42.1	9	2	P82568
2	16	42.1	9	10	Q8LPT5
3	15	39.5	9	9	Q9XJNC
4	14	36.8	9	10	Q9FS22
5	14	36.8	9	11	Q9CZ88
6	14	36.8	9	15	Q85710
7	14	36.8	9	15	Q8UTD7
8	13	34.2	7	8	O99182
9	13	34.2	8	2	P72279
10	13	34.2	8	2	Q51594
11	13	34.2	8	3	Q9HDS4
12	13	34.2	8	9	Q8SBJ0
13	13	34.2	8	9	Q8H9K1
14	13	34.2	9	4	Q9BQT4
15	13	34.2	9	12	Q9ILX6
16	13	34.2	9	13	Q8AUM7

17	12	31.6	7	10	Q49223
18	12	31.6	8	2	Q8KPK4
19	12	31.6	8	2	Q93SR0
20	12	31.6	8	5	Q9N6M5
21	12	31.6	8	6	Q8WNS1
22	12	31.6	8	13	Q9PS69
23	12	31.6	8	15	Q95YK9
24	12	31.6	9	13	Q8AYL5
25	11	28.9	7	2	Q8KMS9
26	11	28.9	7	2	O07J54
27	11	28.9	7	4	Q15897
28	11	28.9	8	3	Q9URB9
29	11	28.9	8	4	Q9P285
30	11	28.9	8	6	Q9BFA7
31	11	28.9	8	7	Q29801
32	11	28.9	9	7	Q29812
33	11	28.9	9	7	Q29797
34	11	28.9	8	8	Q9MSX1
35	11	28.9	8	11	Q8K3Z7
36	11	28.9	9	2	Q47556
37	11	28.9	9	2	Q99193
38	11	28.9	9	4	Q9H4B1
39	10	26.3	7	11	Q6348C
40	10	26.3	7	12	Q66113
41	10	26.3	7	12	Q9VQ10
42	10	26.3	8	2	Q56759
43	10	26.3	8	2	Q9ZIE9
44	10	26.3	8	2	O85406
45	10	26.3	8	2	P83152

#### ALIGNMENTS

RESULT :

P82568 PRELIMINARY; PRT; 9 AA.  
AC P82568;  
DT 0-OCT-2000 (TREMBLrel. 15, Created)  
DT 0-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 0-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE Unknown protein from 2D-page (Fragment).  
CS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
CC Streptococcus.  
CX NCBI TaxID=1314;  
RN 1;  
RP SEQUENCE AND MASS SPECTROMETRY.  
RC STRAIN=GPS4;  
RA Hogan D.A., Du P., Stevenson T.I., Whitton X., Kilby G.W., Rogers J.,  
FA VanBogelen R.A.;  
FT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes  
PT Proteins";  
RL Submitted (MAY-2000) to the SWISS-PROT data bank.  
CC 1- YASS SPECTROMETRY: YN=22592.04; METHOD=ELECTROSPRAY.  
FT NON\_TER 1  
FT NON\_TER 9  
SQ SEQUENCE 9 AA: 1069 MW; 2A771042CB1AB2D7 CRC64;

Query Match 42.1%; Score 16; DB 2; Length 9;  
Best Local Similarity 40.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 DDLJE 8  
Db 4 DEVIE 8

RESULT: 2  
Q8LPT5 PRELIMINARY; PRT; 9 AA.  
ID Q8LPT5  
AC Q8LPT5;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)

Q49223 glycine max  
Q8KPK4 microcystis  
Q93SR0 staphylococ  
Q9N6M5 toxoplasma  
Q8WNS1 bos taurus  
Q9PS69 gallus gall  
Q95YK9 human immun  
Q8AYL5 carassius a  
Q8KMS9 enterobacte  
O07J54 synechococ  
Q15897 homo sapien  
Q9URB9 saccharomyc  
Q9P285 homo sapien  
Q9BFA7 macroscelid  
Q29801 homo sapien  
Q29812 homo sapien  
Q29797 homo sapien  
Q9MSX1 jurinea hum  
Q8K3Z7 mus musculu  
Q47556 escherichia  
Q99193 pseudomonas  
Q9H4B1 homo sapien  
Q6348C rattus norv  
Q66113 cherry leaf  
Q9VQ10 transmissib  
Q56759 xanthobacte  
Q9ZIE9 neisseria m  
O85406 coxiella bu  
P83152 anabaena sp

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Beta-expansin-like protein (Fragment).  
 OS Zea mays (Maize).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 CX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. C123;  
 RA Chang A.S., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,  
 RA Morgante M., Rafalski J.A.;  
 RT "GNP frequency, haplotype structure and linkage disequilibrium in  
 RT elite maize inbred lines";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY094300; AA21836.1; -.  
 FT NON TER 1  
 SQ SEQUENCE 9 AA; 977 MW; 5C05B2D2C81AA3A3 CRC64;

Query Match 42.1%; Score 16; DB 10; Length 9;  
 Best Local Similarity 33.3%; Pred. No. 8.3e+05;  
 Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLEA 9  
 D 4 DEVVDA 9

## RESULT 3

ID Q9XJNO PRELIMINARY; PRT; 9 AA.

AC Q9XJNO;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created);  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update);  
 DT 01-DEC-2001 (TrEMBLrel. 13, Last annotation update);  
 DE P10 (Fragment).  
 CS Bacteriophage phi-10.  
 OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.  
 CX NCBI\_TaxID=90883;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=99350412; PubMed=10419946;  
 RA Mandich L., Qiao X., Qiao J., Onodera S., Romantschuk M.,  
 RA Hoogstraaten D.;  
 RT "Isolation of additional bacteriophages with genomes of segmented  
 RT double-stranded RNA.";  
 RL J. Bacteriol. 181:4505-4508 (1999).  
 DR EMBL; AF125675; AA222555.1; -.  
 FT NON TER 9  
 SQ SEQUENCE 9 AA; 1058 MW; 880E376AA720544A CRC64;

Query Match 39.5%; Score 15; DB 9; Length 9;  
 Best Local Similarity 40.0%; Pred. No. 8.3e+05;  
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
 D 2 DNLID 6

## RESULT 4

ID Q9FSZ2 PRELIMINARY; PRT; 9 AA.

AC Q9FSZ2;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created);  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update);  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update);  
 DE Hypothetical 1.0 kDa protein (Fragment).  
 OS Cicer arietinum (Chickpea) (Garbanzo).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.

CX NCBI\_TaxID=3827;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Castellana; TISSUE=Isolated epicotyl;  
 RA Dopico B., Jimenez T., Labrador E.;  
 RT "cDNA clones expressed in etiolated Cicer arietinum epicotyls.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; A229069; CAC10216.1; -.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 9 AA; 990 MW; 944BDDAA7272EBE CRC64;

Query Match 36.8%; Score 14; DB 10; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLEA 9  
 D 4 LLDA 7

## RESULT 5

ID Q9QZAB PRELIMINARY; PRT; 9 AA.

AC Q9QZAB;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created);  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update);  
 DT 01-DEC-2001 (TrEMBLrel. 13, Last annotation update);  
 DE C-type lectin Dc11 (Fragment).  
 GN Dc11.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Solurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Gotski K., Huang X., Tseng S.-Y., Rattis F., Pardoll D., Tsuchiya H.;  
 RA "Dendritic cell regulation of Dc11 mRNA expression";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF192526; AA204843.1; -.  
 DR XGD; XGI:2136650; Dc11.  
 KW Lectin.  
 FT NON TER 9  
 SQ SEQUENCE 9 AA; 994 MW; 342161AB172EBAB7 CRC64;

Query Match 36.8%; Score 14; DB 10; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 DLJE 8  
 D 3 DLCE 6

## RESULT 6

ID Q85710 PRELIMINARY; PRT; 9 AA.

AC Q85710;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created);  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update);  
 DT 01-NOV-1998 (TrEMBLrel. 05, Last annotation update);  
 DE Pol protein (Fragment).  
 OS Rous sarcoma virus.  
 CC Viruses; Retroviridae; Retroviridae; Alpharetrovirus.  
 CX NCBI\_TaxID=11866;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=84115090; PubMed=6319754;  
 RA Lerner T.L., Hanafusa H.;  
 RT "DNA sequence of the Bryan high-titer strain of Rous sarcoma virus:  
 RT Extent of env deletion and possible genealogical relationship with  
 RT other viral strains";  
 RL J. Virol. 49:549-556 (1984).

```
DR EMBL; K03365; AAA42557.1; -.
FT NON TER 1
SQ SEQUENCE 9 AA; 349 MW; 94AA144DDCD731AA CRC64;

Query Match: 36.8%; Score 14; DB 15; Length 9;
Best Local Similarity 50.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DDLLEA 9
Db 1 EDTLAA 6

RESULT 7
Q8UTD7 PRELIMINARY; PRT; 9 AA.
AC Q8UTD7;
DT C1-MAR-2002 (TrEMBLrel. 20, Created);
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update);
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update);
DE Vpu protein.
GN VPU.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID=111676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=08W1471.27;
RA Novitsky V.A., Smith J.R., Gilbert P., McLane M.F., Chigwedere P.,
RA Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,
RA Foley B.T., Gacikwe S., Rybak N., Gasetiwe S., Vannberg P.,
RA Marlink R., Lee T.-H., Essex M.;
RT "HIV-1 subtype C molecular phylogeny: consensus sequence for an AIDS
RT vaccine design."
RJ Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF443091; AAL347.2.1; -.
SQ SEQUENCE 9 AA; 1102 MW; 188BD40B1727244C CRC64;

Query Match: 36.8%; Score 14; DB 15; Length 9;
Best Local Similarity 75.0%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 DDLLE 8
Db 3 NLLE 6

RESULT 8
O99182 PRELIMINARY; PRT; 7 AA.
AC O99182;
DT C1-MAY-1999 (TrEMBLrel. 10, Created);
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update);
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update);
DE Cytochrome oxidase I (Fragment).
GN COI.
OS Gnatholebias zonatus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulidae; Gnatholebias.
CX NCBI_TaxID=135316;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=10603257;
RA Murphy W.J., Thomson J.E., Collier G.E.;
RA "Phylogeny of the Neotropical killifish family Rivulidae
(Rivulidae) inferred from mitochondrial DNA
sequences."
RJ Mol. Phylogenet. Evol. 13:289-301(1999).
DR EMBL; AF002591; AAD01074.1; -.
KW Mitochondrion.

FT NON TER 1
SQ SEQUENCE 7 AA; 899 MW; 672721F6CB572030 CRC64;

Query Match: 34.2%; Score 13; DB 8; Length 7;
Best Local Similarity 42.9%; Pred. No. 8.3e+05;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLXDDL 7
Db 1 ILYQHLL 7

RESULT 9
P72279 PRELIMINARY; PRT; 8 AA.
AC P72279;
DT 01-FEB-1997 (TrEMBLrel. 02, Created);
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update);
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update);
DE Biphényl dioxxygenase (Fragment).
GN BPHB.
OS Rhodococcus globerulus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
CX NCBI_TaxID=33038;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95255652; PubMed=7737502;
RA Asturias J.A., Diaz E., Timmis K.N.;
RA "Evolutionary relationship of the biphenyl dioxxygenase of the gram-
RA positive bacterium Rhodococcus globerulus p6 to multicomponent
RA dioxxygenases of gram-negative bacteria."
RJ J. Bacteriol. 166:1106-1112(1986).
DR EMBL; X80341; CAA56350.1; -.
KW Dioxxygenase.
FT NON TER 8
SQ SEQUENCE 8 AA; 989 MW; EBD2CBIAB6D73406 CRC64;

Query Match: 34.2%; Score 13; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 8.3e+05;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDDL 7
Db 3 LQDEVV 8

RESULT 10
Q51594 PRELIMINARY; PRT; 8 AA.
AC Q51594;
DT 01-NOV-1996 (TrEMBLrel. 01, Created);
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update);
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update);
DE CopB protein (Fragment).
OS Escherichia coli.
CX Plasmid ColV2-K94.
GN ColV2-K94.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=86223772; PubMed=2423502;
RA Weber P.C., Palchaudhuri S.;
RA "Incompatibility repressor in a repA-like replicon of the IncFI
RA plasmid ColV2-K94."
RJ J. Bacteriol. 166:1106-1112(1986).
DR EMBL; M13472; AAA23194.1; -.
KW Plasmid.
FT NON TER 1
SQ SEQUENCE 8 AA; 998 MW; 0F37205AA73416D7 CRC64;

Query Match: 34.2%; Score 13; DB 2; Length 8;
```

Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLE 8  
DB 5 DLK 8

RESULT 11  
Q9HDS4 PRELIMINARY; PRT; 8 AA.  
AC Q9HDS4  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE TrpC polypeptide (Fragment).  
GN TRPC.  
OS Aspergillus flavus.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
OX NCBI\_TaxID=5059;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A55;  
RA Geiser J.W., Dorner J.W., Horn B.W., Taylor J.W.;  
RT "The phylogenetics of mycotoxin and sclerotium production in  
RT Aspergillus flavus and Aspergillus oryzae."  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF261861; AAG16135.1; --  
KW Polypeptide.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 807 MW; F3B2C72A95B87DD6 CRC64;

Query Match 34.2%; Score 13; DB 3; Length 8;  
Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLE 8  
DB 5 DLVD 8

RESULT 12  
Q8SBJ0 PRELIMINARY; PRT; 8 AA.  
AC Q8SBJ0  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Gp30.2 (Fragment).  
GN 30.2.  
OS Bacteriophage RB69.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC T4-like viruses.  
OX NCBI\_TaxID=12353;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Plesinskiene G., Kolesinskiene G., Truncaite L., Zajackauskaite A.,  
RA Nivinskas R.,  
RT "Genomic region with genes 30.6-30.3 of T4-related bacteriophages."  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ439452; CAD28423.1; --  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 918 MW; F3B1A72041B76336 CRC64;

Query Match 34.2%; Score 13; DB 9; Length 8;  
Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXD 4  
DB 5 ILTD 8

RESULT 13  
Q8H9K1 PRELIMINARY; PRT; 8 AA.  
AC Q8H9K1  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Gp30.2 protein (Fragment).  
GN 30.2.  
OS Bacteriophage L210.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC T4-like viruses.  
OX NCBI\_TaxID=192373;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kolesinskiene G., Nivinskas R.,  
RT "A pair of overlapping genes 30.3 and 30.3' of T4-related  
RT bacteriophages."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ458400; CAD30256.1; --  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 918 MW; F3B1A72041B76336 CRC64;

Query Match 34.2%; Score 13; DB 9; Length 8;  
Best Local Similarity 50.0%; Pred. No. 8.3e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXD 4  
DB 5 ILTD 8

RESULT 14  
Q9BQT4 PRELIMINARY; PRT; 9 AA.  
AC Q9BQT4  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CLIP-associating protein CLASP2 (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Gajart N.,  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DE Gp30.2 (Fragment).  
GN 30.2.  
OS Bacteriophage RB69.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC T4-like viruses.  
OX NCBI\_TaxID=12353;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Plesinskiene G., Kolesinskiene G., Truncaite L., Zajackauskaite A.,  
RA Nivinskas R.,  
RT "Genomic region with genes 30.6-30.3 of T4-related bacteriophages."  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ439452; CAD28423.1; --  
FT NON\_TER 8  
SQ SEQUENCE 9 AA; 1001 MW; C35CB33AAA8769D CRC64;

Query Match 34.2%; Score 13; DB 4; Length 9;  
Best Local Similarity 50.0%; Pred. No. 6.3e+05;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDD 5  
DB 3 MGDD 6



```

RESULT 15
Q9ILX6
ID Q9ILX6 PRELIMINARY; PRI; 9 AA.
AC Q9ILX6;
CT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DI 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ORF 10-like protein (Fragment).
OS Macaca nemestrina rhadinovirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI TaxID=123630;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=Mne42N;
RX MEDLINE=20240083; PubMed=10775636;
RA Schultz E.R., Rankin G.W. Jr., Blanc M.P., Raden B.W., Tsai C.C.,
RA Rose T.M.;
RT "Characterization of two divergent lineages of macaque rhadinoviruses
RT related to Kaposi's sarcoma-associated herpesvirus.";
RL J. Virol. 74:4919-4928(2000).
DR EMBL; AF204167; AAF81665.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1003 MW; 415BD587233AB2C7 CRC64;

Query Match 34.2%; Score 13; DB 12; Length 9;
Best Local Similarity 50.0%; Pred No. 8.3e-05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CY 1 VLXDJL 6
DB 1 MLVDKL 6

```

Search completed: October 20, 2003, 13:15:50  
Job time : 97 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

CM: protein - protein search, using sw model

Run on: October 20, 2003, 13:03:24 ; Search time 81 Seconds

(without alignments)

17.636 Million cell updates/sec

Title: US-09-489-760A-1

Perfect score: 32

Sequence: 1 VXXDLLLEA 9

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1:07863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 179625

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jur01:\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-980.DAT:\*

2: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-981.DAT:\*

3: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-982.DAT:\*

4: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-983.DAT:\*

5: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-984.DAT:\*

6: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-985.DAT:\*

7: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-986.DAT:\*

8: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-987.DAT:\*

9: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-988.DAT:\*

10: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-989.DAT:\*

11: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-990.DAT:\*

12: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-991.DAT:\*

13: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-992.DAT:\*

14: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-993.DAT:\*

15: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-994.DAT:\*

16: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-995.DAT:\*

17: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-996.DAT:\*

18: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-997.DAT:\*

19: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-998.DAT:\*

20: /SIDSI/gcgdata/geneseq/geneseq-emb/AA-999.DAT:\*

21: /SIDSI/gcgdata/geneseq/geneseq-emb/AA2000.DAT:\*

22: /SIDSI/gcgdata/geneseq/geneseq-emb/AA2001.DAT:\*

23: /SIDSI/gcgdata/geneseq/geneseq-emb/AA2002.DAT:\*

24: /SIDSI/gcgdata/geneseq/geneseq-emb/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	94.7	9	20	AAW97374 HA-1 R-allele sequ
2	36	94.7	9	20	AAW97375 HA-1 H-allele sequ
3	36	94.7	9	20	AAW97572 T-cell epitope fro
4	36	94.7	9	20	AAW99195 Minor histocompati
5	36	94.7	9	20	AAW99196 Minor histocompati
6	36	94.7	9	20	AAW99197 Minor histocompati
7	23	60.5	9	23	AAU71211 Human MHC class I
8	23	60.5	9	23	AAU71428 Human MHC molecule
9	22	57.9	9	21	AAW90162 HLA-A2 restricted

10	22	57.9	9	23	AAJ03449 Chlamydia pneumonia
11	22	57.9	9	24	ABP75296 Proteome analysis
12	22	57.9	9	24	ABP75138 Proteome analysis
13	22	57.9	9	24	ABP75181 Peptide # 2 used i
14	21	55.3	6	21	AAI12035 Rat Rb-interacting
15	21	55.3	6	21	AAU80810 Ad7 cell peptide
16	21	55.3	8	21	AAI12089 Prenyl diphosphate
17	21	55.3	9	19	AAW47427 Peptide epitope of
18	21	55.3	9	20	AAW97373 Cell motif peptide
19	21	55.3	9	21	AAI12086 Ad43 cell peptide
20	21	55.3	9	21	AAI12092 Human C35 peptide
21	21	55.3	9	22	ABG77888 Human C35 peptide
22	21	55.3	9	22	ABG12498 Human C35 peptide
23	21	55.3	9	22	ABG12506 Human C35 peptide
24	21	55.3	9	22	ABG12554 Human C35 peptide
25	21	55.3	9	22	ABG12579 Human C35 peptide
26	21	55.3	9	22	ABG12651 Human C35 peptide
27	21	55.3	9	22	ABG12658 Human C35 peptide
28	21	55.3	9	22	ABG12737 Human C35 peptide
29	21	55.3	9	22	ABG12755 Human C35 peptide
30	21	55.3	9	22	ABG12774 Human C35 peptide
31	21	55.3	9	22	ABG12808 Human C35 peptide
32	21	55.3	9	22	ABG12849 Human C35 peptide
33	21	55.3	9	22	ABG12983 Human C35 peptide
34	21	55.3	9	22	ABG12995 Human C35 peptide
35	21	55.3	9	22	ABG13032 Human C35 peptide
36	21	55.3	9	22	ABG13072 Human C35 peptide
37	21	55.3	9	22	ABG13115 Human C35 peptide
38	21	55.3	9	22	ABG13144 Human C35 peptide
39	21	55.3	9	22	ABG13160 Human C35 peptide
40	21	55.3	9	22	ABG13184 Human C35 peptide
41	21	55.3	9	22	ABG13186 Human C35 peptide
42	21	55.3	9	22	ABG13204 Human C35 peptide
43	21	55.3	9	22	ABG13206 Human C35 peptide
44	21	55.3	9	22	ABG13226 Human C35 peptide
45	21	55.3	9	22	ABG13228 Human C35 peptide

#### ALIGNMENTS

RESULT :

AAW97374

ID AAW97374 standard; Protein: 9 AA.

AC AAW97374;

XX

XX

DT 13-YAY-1999 (first entry)

XX

XX

DE HA-1 R-allele sequence.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT Typing minor histocompatibility antigen HA-1 - by amplifying and  
 PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g.  
 PT detection of genetic aberrances

XX Claim 13; Fig 5; 59pp; English.

XX The present sequence represents part of the minor histocompatibility  
 CC antigen HA-1 R-allele. The specification describes methods for typing  
 CC alleles (preferably the H and R alleles) of the minor histocompatibility  
 CC antigen HA-1 in a sample, which comprise detecting polymorphic  
 CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The  
 CC methods can be used for HA-1 typing for bone marrow transplants, severe  
 CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as  
 CC detection of genetic aberrances. The probes and primers of the invention  
 CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be  
 CC used anti-idiotypic B cells and/or T cells and antibodies.

XX Sequence 9 AA;

Query Match 94.7%; Score 36; DB 20; Length 9;  
 Best Local Similarity 88.9%; Pred.No. 9.3e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9

DB 1 VLXDDLEA 9

RESULT 2

AAW97375  
 ID AAW97375 standard; Protein; 9 AA.

AC AAW97375;

DT 13-MAY-1999 (first entry)

DE HA-1 H-allele sequence.

XX Introns; minor histocompatibility antigen HA-1; typing allele;  
 XX H allele; R allele; polymorphic nucleotide; HA-1 typing;  
 XX bone marrow transplant; severe aplastic anaemia; leukaemia;  
 XX immune deficiency disease; ss.

XX Homo sapiens.

XX WO9905313-A2.

XX 04-FEB-1999.

XX 23-JUL-1998; 98WO-EP04928.

XX 02-JUN-1998; 98EP-0870125.

XX 23-JUL-1997; 97EP-0202303.

XX (UYLE-) RIJCSUNIV LEIDEN.

XX Goulmy E;

XX WPI; 1999-142960/12.

PT Typing minor histocompatibility antigen HA-1 - by amplifying and  
 PT identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g.  
 PT detection of genetic aberrances

XX Claim 18; Fig 5; 59pp; English.

XX The present sequence represents part of the minor histocompatibility  
 CC antigen HA-1 H-allele. The specification describes methods for typing  
 CC alleles (preferably the H and R alleles) of the minor histocompatibility  
 CC antigen HA-1 in a sample, which comprise detecting polymorphic  
 CC nucleotides in the cDNA or genomic nucleic acids of the alleles. The  
 CC methods can be used for HA-1 typing for bone marrow transplants, severe  
 CC aplastic anaemia, leukaemia and immune deficiency diseases, as well as

CC detection of genetic aberrances. The probes and primers of the invention  
 CC can be used to screen for the HA-1 alleles. The HA-1 peptides can be  
 CC used anti-idiotypic B cells and/or T cells and antibodies.

XX Sequence 9 AA;

Query Match 94.7%; Score 36; DB 20; Length 9;  
 Best Local Similarity 88.9%; Pred.No. 9.3e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9

DB 1 VLXDDLEA 9

RESULT 3

AAW97572

ID AAW97572 standard; peptide; 9 AA.

XX AAW97572;

XX 20-MAY-1999 (first entry)

DE T-cell epitope from the minor histocompatibility antigen HA-1.

XX T-cell epitope; minor histocompatibility antigen HA-1; vaccine;  
 XX transplant rejection; Graft-versus-Host Disease; autoimmune disease;  
 XX neoplastic haematopoietic cell.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 3

XX /note= "His or Arg"

XX WO9905173-A1.

XX 04-FEB-1999.

XX 23-JUL-1998; 98WO-NL00424.

XX 23-JUL-1997; 97EP-0202303.

XX (UYLE-) RIJCSUNIV LEIDEN.

XX Engelhard VE, Goulmy ECM, Hunt DF;

XX WPI; 1999-142855/12.

PT Immunogenic peptide from minor histocompatibility antigen HA-1 -  
 PT useful for inducing tolerance to transplants and prevent rejection  
 PT or graft-versus-host disease

XX Claim 1; Page 39; 57pp; English.

XX The present sequence represents an immunogenic peptide constituting a  
 CC T-cell epitope, obtainable from the minor histocompatibility antigen  
 CC HA-1. The peptide can be used in vaccines or pharmaceutical formulations  
 CC as medicines to induce tolerance for transplants so as to prevent  
 CC rejection and/or Graft-versus-Host Disease, or to treat autoimmune  
 CC diseases. Neoplastic haematopoietic cells presenting the peptides, in  
 CC an HLA class I context, can be eliminated after specific recognition  
 CC of the peptides. The peptides can also be used to raise antibodies,  
 CC T-cell receptor, B- and T-cells.

XX Sequence 9 AA;

Query Match 94.7%; Score 36; DB 20; Length 9;

Best Local Similarity 100.0%; Pred.No. 9.3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9

DB 1 VLXDDLEA 9

```

DB      1 VLXDDLEA 9
RESULT 4
ID      AAW99195 standard; peptide; 9 AA.
XX      AC      AAW99195;
XX      DT      20-MAY-1999 (first entry);
XX      DE      Minor histocompatibility antigen HA-1 T-cell epitope #1.
XX      KW      Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
XX      KW      graft versus host disease; bone marrow transplant; leukaemia; vaccine;
XX      KW      diagnosis; aplastic anaemia; immune deficiency disease.
XX      OS      Homo sapiens.
XX      OS      Synthetic.
XX      PH      Key      Location/Qualifiers
XX      FT      Misc-difference 3
XX      FT      /label= His, Arg
XX      PN      WO9905174-A1.
XX      PD      04-FEB-1999.
XX      PF      23-JUL-1998; 98WO-NL00425.
XX      PR      23-JUL-1997; 97EP-0202303.
XX      PA      (UYLE-) RIJKSUNIV LEIDEN.
XX      PI      Engelhard VH, Goulmy EAJM, Hunt DF;
XX      DR      WPI; 1999-153312/13.
XX      PT      A new minor histocompatibility antigen, HA-1 - useful to treat
XX      PT      immune diseases and prevent rejection and host versus graft disease
XX      PT      in bone marrow and organ transplantation
XX      PS      Claim 1; Page 32; 47pp; English.
XX      CC      The present sequence represents a new peptide (P1) constituting a T-cell
XX      CC      epitope obtainable from the minor histocompatibility antigen HA-1. The
XX      CC      peptide is immunogenic and can be used as part of a vaccine. P1 is used
XX      CC      as a medicine, to induce tolerance for transplants, prevent rejection
XX      CC      and/or graft versus host disease, or to treat (auto) immune diseases.
XX      CC      In particular it can be used with bone marrow transplantation, in the
XX      CC      treatment of severe aplastic anaemia, leukaemia, and immune deficiency
XX      CC      diseases.
XX      SQ      Sequence 9 AA;
XX      Query Match      94.7%; Score 36; DB 20; Length 9;
XX      Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX      Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX      QY      1 VLXDDLEA 9
XX      DB      1 VLXDDLEA 9
XX      RESULT 6
XX      ID      AAW99197 standard; peptide; 9 AA.
XX      AC      AAW99197;
XX      DT      20-MAY-1999 (first entry);
XX      DE      Minor histocompatibility antigen HA-1 T-cell epitope #3.
XX      KW      Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;
XX      KW      graft versus host disease; bone marrow transplant; leukaemia; vaccine;
XX      KW      diagnosis; aplastic anaemia; immune deficiency disease.
XX      OS      Homo sapiens.
XX      OS      WO9905174-A1.
XX      PD      04-FEB-1999.
XX      PF      23-JUL-1998; 98WO-NL00425.
XX      PR      23-JUL-1997; 97EP-0202303.
XX      PA      (UYLE-) RIJKSUNIV LEIDEN.
XX      PI      Engelhard VH, Goulmy EAJM, Hunt DF;
XX      DR      WPI; 1999-153312/13.
XX      PT      A new minor histocompatibility antigen, HA-1 - useful to treat
XX      PT      immune diseases and prevent rejection and host versus graft disease
XX      PT      in bone marrow and organ transplantation
XX      PS      Claim 1; Page 32; 47pp; English.
XX      CC      The present sequence represents a new peptide (P1) constituting a T-cell
XX      CC      epitope obtainable from the minor histocompatibility antigen HA-1. The
XX      CC      peptide is immunogenic and can be used as part of a vaccine. P1 is used
XX      CC      as a medicine, to induce tolerance for transplants, prevent rejection
XX      CC      and/or graft versus host disease, or to treat (auto) immune diseases.
XX      CC      In particular it can be used with bone marrow transplantation, in the
XX      CC      treatment of severe aplastic anaemia, leukaemia, and immune deficiency
XX      CC      diseases.
XX      SQ      Sequence 9 AA;
XX      Query Match      94.7%; Score 36; DB 20; Length 9;
XX      Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX      Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX      QY      1 VLXDDLEA 9
XX      DB      1 VLXDDLEA 9
XX      RESULT 5
XX      ID      AAW99196 standard; peptide; 9 AA.
XX      AC      AAW99196;
XX      DT      20-MAY-1999 (first entry);
XX      DE      Minor histocompatibility antigen HA-1 T-cell epitope #2.

```

```

PI Engelhard VH, Goulmy EACM, Hunt DF;
XX WPI; 1999-153312/13.
XX
XX A new minor histocompatibility antigen, HA-1 - useful to treat
XX immune diseases and prevent rejection and host versus graft disease
XX in bone marrow and organ transplantation
XX
XX Disclosure; Page 15; 47pp; English.
XX
XX The present sequence represents a new peptide (PI) constituting a T-cell
XX epitope obtainable from the minor histocompatibility antigen HA-1. The
XX peptide is immunogenic and can be used as part of a vaccine. PI is used
XX as a medicine, to induce tolerance for transplants, prevent rejection
XX and/or graft versus host disease, or to treat (auto) immune diseases.
XX In particular it can be used with bone marrow transplantation, in the
XX treatment of severe aplastic anaemia, leukaemia, and immune deficiency
XX diseases.
XX
XX Sequence 9 AA;
XX
XX Query Match 94.7%; Score 36; DB 20; Length 9;
XX Best Local Similarity 88.9%; Pred. No. 9.3e+05;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 VLXDDLEA 9
XX DB 1 VLXDDLEA 9
XX
XX RESULT 7
XX AAU7121;
XX ID AAU71211 standard; Peptide; 9 AA.
XX
XX AC AAU71211;
XX
XX DT 26-FEB-2002 (first entry)
XX
XX DE Human MHC class I molecule HLA-A2 binding 103P3E8 peptide #16.
XX
XX KW 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;
XX tumor; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;
XX single chain monoclonal antibody; serum; blood; urine; tissue; human;
XX chromosome 9q13-q21.
XX
XX OS Homo sapiens.
XX
XX FN WO200179557-A2.
XX
XX FD 25-OCT-2001.
XX
XX PF 12-APR-2001; 2001WO-US12181.
XX
XX PR 12-APR-2000; 2000US-196647P.
XX
XX PA (UROC-); UROGENESYS INC.
XX
XX PI Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;
XX PI Jakobovits A;
XX
XX DR WPI; 2002-061976/08.
XX
XX PS Disclosure; Page 84; 128pp; English.
XX
XX PT Monitoring 103P3E8 gene products in sample from patient (suspected of)
XX having cancer, useful for diagnosing, managing or treating cancers,
XX e.g. prostate cancer, comprises determining presence of aberrant
XX 103P3E8 gene products -
XX
XX PS Disclosure; Page 84; 128pp; English.
XX
XX Sequences AAU71093-AAU7196 represent the 103P3E8-related protein and
XX peptide fragments of the protein. 103P3E8 exhibits tissue specific
XX expression in normal adult tissue, but it is also aberrantly expressed in
XX many cancers including tumours of the prostate, bladder, kidney, colon,
XX
XX PT Monitoring 103P3E8 gene products in sample from patient (suspected of)
XX having cancer, useful for diagnosing, managing or treating cancers,
XX e.g. prostate cancer, comprises determining presence of aberrant
XX 103P3E8 gene products -
XX
XX PS Disclosure; Page 84; 128pp; English.
XX
XX Sequences AAU71093-AAU7196 represent the 103P3E8-related protein and
XX peptide fragments of the protein. 103P3E8 exhibits tissue specific
XX expression in normal adult tissue, but it is also aberrantly expressed in
XX many cancers including tumours of the prostate, bladder, kidney, colon,
XX
XX lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related
XX protein and peptide fragments and specific PCR primers are therefore
XX useful for diagnosing and treating cancer. A vector comprising a
XX polynucleotide which encodes a single chain monoclonal antibody, that
XX immunospecifically binds to an 103P3E8-related protein, and a ribozyme
XX capable of cleaving a polynucleotide having the 103P3E8 coding sequence,
XX are both useful in the preparation of a composition for treating a

```

CC patient with a cancer that expresses 103P3E8. The sequences can be used  
 CC in diagnostic methods to monitor the level of 103P3E8 gene products in:  
 CC serum, blood, urine and tissue and to thereby detect the presence of  
 CC cancerous cells.

XX Sequence 9 AA;

Query Match 60.5%; Score 23; DB 03; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLEA 9

DB 5 DLLEA 9

RESULT 9

ID AAY80162  
 AC AAY80162; peptide; 9 AA.

XX 24-MAY-2000 (first entry)

XX HLA-A2 restricted tumour antigen peptide derived from SART-1 #15.  
 DE HLA-A2 restricted tumour antigen; SART-1; human leukocyte antigen;  
 XX human; diagnosis; tumour; cytotoxic T cell; flat epithelioma;  
 KW lung cancer; oesophagus cancer; cytostatic.

XX Homo sapiens.

CS Synthetic.

XX WO200006595-A1.

XX 10-FEB-2000.

XX 27-JUL-1999; 99WC-JP040:0.

XX 28-JUL-1998; 98JP-0212940.

XX (SJMU) SUMITOMO PHARM CO LTD.

XX (ITOHI) ITOH K.

XX Itoh K, Kobayashi T;

XX MPI; 2000-195258/17.

XX HLA-A2 resistant tumor antigen peptides derived from SART-1, useful as  
 PT preventives or diagnostics for tumors e.g. flat epithelioma like lung  
 PT cancer.

PS Claim 2; Page 41; 55pp; Japanese.

XX AAY80148 to AAY80181 represent specifically claimed HLA-A2 restricted  
 CC tumour antigen peptides derived from SART-1. The peptides have  
 CC cytostatic activity. The peptides are useful as a preventive or  
 CC diagnostic for tumors e.g. flat epithelioma like lung cancer and  
 CC oesophagus cancer.

XX Sequence 9 AA;

Query Match 57.9%; Score 22; DB 21; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VJXDDL 7

DB 3 VRADDL 9

RESULT 10

AAY80149

AAU09449 standard; peptide; 9 AA.

XX AAU09449;

XX 26-MAR-2002 (first entry)

XX Chlamydomydia pneumoniae T-cell epitope #4 useful as Chlamydia antigen.

XX ATP binding cassette; secretory locus open reading frame; endopeptidase;  
 KW secretory locus ORF; protease; metalloprotease; CLP protease; Arpase;  
 KW CLP protease subunit; transglycolase/transpeptidase; CLP protease;  
 KW thioesterin; Chlamydia infection; antibacterial; immunogen.

XX Chlamydomydia pneumoniae CWL029.

XX WO2000:85972-A2.

XX 15-NOV-2001.

XX 08-MAY-2001; 2001WO-CA00653.

XX 08-MAY-2000; 2000US-202672P.

XX 30-MAY-2000; 2000US-207852P.

XX 16-JUN-2000; 2000US-211796P.

XX 16-JUN-2000; 2000US-211797P.

XX 16-JUN-2000; 2000US-211798P.

XX 16-JUN-2000; 2000US-211801P.

XX 16-JUN-2000; 2000US-212044P.

XX 26-SEP-2000; 2000US-23535P.

XX 26-SEP-2000; 2000US-23536P.

XX (AVET) AVENTIS PASTEUR LTD.

XX Murdin AD, Oomen RP, Wang J, Dunn P;

XX WP; 2002-04947/06.

XX Vaccine useful for immunising mammals against chlamydia infections.

XX PT comprises vectors having sequences of Arp binding cassette gene,

XX secretory locus open reading frame gene of chlamydia.

XX Example 4; Page 80; 355pp; English.

XX The present invention relates to the isolation of Chlamydomydia  
 CC pneumoniae strain CWL029 genes and their encoded proteins. The genes of  
 CC the invention encode an Arp binding cassette gene, a secretory locus  
 CC open reading frame (ORF), an endopeptidase, a protease, a  
 CC metalloprotease, CLP protease Arpase, a CLP protease subunit, a  
 CC transglycolase/transpeptidase, a CLP protease, or thioesterin. The  
 CC genes of the invention can be used in a vector as a vaccine for the  
 CC prevention and treatment of Chlamydia infections. AAU09448-AAU09473  
 CC represent B- or T-cell epitopes from the C. pneumoniae proteins  
 CC (AAU09430-AAU09439) of the invention. These epitopes can be used as  
 CC Chlamydia antigens.

XX Sequence 9 AA;

Query Match 57.9%; Score 22; DB 23; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLLLE 8

DB 2 LGDBILE 8

RESULT 11

ABP75298

ID ABP75298 standard; Peptide; 9 AA.

XX ABP75298;

XX AC ABP75298;

DT 20-FEB-2003 (first entry)  
DE Chlamydia pneumonia peptide epitope #64.  
XX Antibacterial; secreted protein; intracellular bacterium.  
XX Chlamydia pneumonia.  
XX WO200282091-A2.  
XX 17-OCT-2002.  
XX 09-APR-2002; 2002WO-DK00234.  
XX 09-APR-2001; 2001DK-0000581.  
XX 09-APR-2001; 2001US-282513P.  
XX (SHAW/) SHAW A C.  
XX (VAND/) VANDAHN B B.  
XX Shaw AC, Vandahl BB;  
XX WPI; 2003-058565/05.  
XX Identifying intracellular bacterial proteins by labeling proteins in  
PT the presence of a eukaryotic protein synthesis inhibitor, performing  
PT electrophoresis, autoradiography and comparing profiles to an  
PT infected-cell lysate profile -  
XX Claim 34; Page 151; 179pp; English.  
XX The present invention relates to a method (M1) for identifying secreted  
CC intracellular bacterial proteins (BP); M1 comprises: (a) selectively  
CC visualizing BP by pulse labelling in the presence of an inhibitor of  
CC eukaryotic protein synthesis followed by 2D electrophoresis and  
CC autoradiography; (b) comparing protein profiles (BP) of purified bacteria  
CC to PF of total lysate (TL) of infected cells; and (c) identifying protein  
CC spots present in differential images from gels loaded with TL. The  
CC present sequence is one such bacterial peptide epitope which was  
CC identified by the method of the invention.  
XX Sequence 9 AA;  
SQ  
Query Match 57.9%; Score 22; DB 24; Length 9;  
Best Local Similarity 57.1%; Pred. No. 9.3e+05;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2 LKDDLE 8  
DB 2 LGDELE 9  
RESULT 12  
ABP75138  
XX ABP75138 standard; Peptide; 9 AA.  
XX AC ABP75138;  
XX DT 03-FEB-2003 (first entry)  
XX Proteome analysis related peptide #423.  
XX Proteome analysis; isolation; determination; diagnostic assay; detection;  
KW protein marker; identification; metastatic; invasive cancer;  
KW differential expression; signalling pathway; chromatography.  
XX Synthetic.  
XX WC200277016-A2.  
XX 03-OCT-2002.  
XX 22-MAR-2002; 2002WC-EP03368.  
XX 22-MAR-2001; 2001US-278171P.  
XX 12-SEP-2001; 2001US-318749P.  
XX 20-SEP-2001; 2001US-323999P.  
XX (VLAAS) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
XX Vandekerckhove C, Gevaert K;  
XX WPI; 2003-067379/06.  
XX Method for isolation of peptides from complex mixture of peptides  
PT involves specific chemical and/or enzymatic alteration of at least one  
PT type of peptide -  
XX Example 22; Page 150; 193pp; English.  
XX The present invention describes a method (M1) for the isolation of a  
CC subset of peptides from a protein peptide mixture (PI). M1 involves:  
CC (a) separating the protein peptide mixture into fractions of peptides via  
CC chromatography; (b) chemically, or enzymatically, or chemically and  
CC enzymatically, altering at least one amino acid of at least one of the  
CC peptides in each fraction, thereby generating a subset of altered  
CC peptides; and (c) isolating the altered (flagged) peptides out of each  
CC fraction via chromatography, where the chromatography of steps (a) and  
CC (c) is performed with the same type of chromatography. M1 can be used for  
CC the isolation and determination of peptides from protein peptide  
CC mixtures. M1 can also be used in diagnostic assays for detection of the  
CC presence, the absence or a variation in expression level of at least one  
CC protein marker or a specific set of proteins indicative of a disease  
CC state. M1 can be used for identifying target proteins present in  
CC metastatic and invasive cancers, in differential expression of proteins  
CC in transgenic mice, identification of proteins that are upregulated or  
CC down regulated in disease tissues, in identification of intracellular  
CC changes in cells with physiological changes such as metabolic shift, in  
CC the identification of biomarkers in cancers and in the identification  
CC of signalling pathways. The method is gel-free methodology for  
CC qualitative and quantitative proteome analysis without the need for  
CC multidimensional chromatography and without the use of affinity tags.  
CC ABP74714 to ABP75190 represent peptide sequences used in the  
CC exemplification of the present invention.  
XX Sequence 9 AA;  
SQ  
Query Match 57.9%; Score 22; DB 24; Length 9;  
Best Local Similarity 71.4%; Pred. No. 9.3e+05;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LKDDLE 8  
DB 3 LGDELE 9  
RESULT 13  
ABP75182  
XX ABP75181 standard; Peptide; 9 AA.  
XX AC ABP75182;  
XX DT 03-FEB-2003 (first entry)  
XX Proteome analysis related peptide #466.  
XX Proteome analysis; isolation; determination; diagnostic assay; detection;  
KW protein marker; identification; metastatic; invasive cancer;  
KW differential expression; signalling pathway; chromatography.  
XX Synthetic.  
XX WC200277016-A2.  
XX 03-OCT-2002.  
XX 03-OCT-2002.

```

PF 22-MAR-2002; 2002WC-EP03368.
XX
XX 22-MAR-2002; 2002US-27871P.
PR 12-SEP-2001; 2001US-318749P.
PR 20-SEP-2001; 2001US-323999P.
XX
XX (VJAA-) VLAAY'S INTERUNIVERSITAIR INST BIOTECHNOS.
PA
PA Vandeckerckhove J, Gevaert K;
PI
XX WP1; 2003-367379/36.
DR
XX
XX Method for isolation of peptides from complex mixture of peptides
PT involves specific chemical and/or enzymatic alteration of at least one
PT type of peptide
XX
XX Example 22; Page 151; 193pp; English.
PS
XX The present invention describes a method (M1) for the isolation of a
CC subset of peptides from a protein peptide mixture (P1). M1 involves:
CC (a) separating the protein peptide mixture into fractions of peptides via
CC chromatography; (b) chemically, or enzymatically, or chemically and
CC enzymatically, altering at least one amino acid of at least one of the
CC peptides in each fraction, thereby generating a subset of altered
CC peptides; and (c) isolating the altered (flagged) peptides out of each
CC fraction via chromatography, where the chromatography of steps (a) and
CC (c) is performed with the same type of chromatography. M1 can be used for
CC the isolation and determination of peptides from protein peptide
CC mixtures. M1 can also be used in diagnostic assays for detection of the
CC presence, the absence or a variation in expression level of at least one
CC protein marker or a specific set of proteins indicative of a disease
CC state. M1 can be used for identifying target proteins present in
CC metastatic and invasive cancers, in differential expression of proteins
CC in transgenic mice, identification of proteins that are upregulated or
CC down regulated in disease tissues, in identification of intracellular
CC changes in cells with physiological changes such as metabolic shift, in
CC the identification of biomarkers in cancers and in the identification
CC of signalling pathways. The method is gel-free methodology for
CC qualitative and quantitative proteome analysis without the need for
CC multidimensional chromatography and without the use of affinity tags.
CC ABF74714 to ABP75193 represent peptide sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 9 AA;
SQ
Query Match 57.9%; Score 22; DB 24; Length 9;
Best Local Similarity 71.4%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 2 LKDDLLS 8
DB 3 LKDDLS 9
RESULT 14
AAB-2035
ID AAB12035 standard; Peptide; 6 AA.
XX
XX AAB12035;
XX
XX 17-JAN-2001 (first entry)
XX
XX Peptide # 2 used in PR domain peptide analysis.
XX
XX Rb-interacting zinc finger; RIZ; retinoblastoma; heart disease;
XX cell proliferation; cell differentiation; tissue repair; PR domain;
XX transcription regulator; breast cancer; gene therapy; melanoma;
XX neuroblastoma; leukaemia; Parkinson's disease; Huntington's disease;
XX Alzheimer's disease; paralysis; motor neuron disorder.
XX
XX Unidentified.
XX
XX US6069231-A.
PD

```

---

```

XX 30-MAY-2000.
XX
XX 18-AUG-1995; 95US-0516859.
XX
XX 18-AUG-1994; 94US-0292683.
XX 06-MAR-1995; 95US-0399411.
XX
XX (LJOL-) LA JOLLA CANCER RES FOUND.
XX
XX Huang S;
XX
XX WP1; 2000-410879/35.
XX
XX New PR domain peptides comprising amino acid sequences from, for
PT example retinoblastoma-interacting zinc finger, or egi-43 proteins, for
PT regulating gene transcription and controlling cell proliferation and
PT differentiation
XX
XX Disclosure; Columns 87-88; 91pp; English.
XX
XX Retinoblastoma (Rb)-interacting zinc finger (RIZ) protein is a nuclear
CC phosphoprotein that acts as a cell differentiation factor. RIZ can
CC modulate cell growth by binding to Rb protein, which is involved in
CC regulating cell proliferation. In addition, RIZ can act to regulate
CC transcription. RIZ functions to maintain cells in the G1 phase of the
CC cell cycle, by interacting with Rb through the cr2 domain of RIZ. RIZ
CC protein is a PR domain protein and is present primarily in the cell
CC nucleus. RIZ gene mutations may be implicated in various cancers such as
CC melanoma, neuroblastoma, leukaemia and breast cancer, and so the RIZ gene
CC may be used in gene therapy for these disorders. Since RIZ protein is
CC implicated in cell cycle arrest, inhibition of RIZ activity may be useful
CC in neurodegenerative disorder therapy e.g. for Parkinson's, Huntington's
CC or Alzheimer's disease, paralysis or motor neuron disorders, or cardiac
CC disorders e.g. heart disease, where the ability to induce neural/cardiac
CC tissue proliferation would be useful. The present sequence is a peptide
CC used in the analysis of PR domain peptides of the present invention.
XX
XX Sequence 6 AA;
SQ
Query Match 55.3%; Score 21; DB 21; Length 6;
Best Local Similarity 80.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CY 4 DDLLE 8
DB 1 EDLLE 5
RESULT 15
AAB-2035
ID AAB12035 standard; Peptide; 6 AA.
XX
XX AAB12035;
XX
XX 26-MAR-2002 (first entry)
XX
XX Rat Rb-interacting zinc finger (RIZ) E1A related motif #2.
XX
XX Rat; human; retinoblastoma protein-interacting zinc finger protein; RIZ;
XX PRD1-BF1; human positive regulatory domain 1-binding factor 1; EVI-1;
XX human ecotropic viral integration site-1 myeloid transforming gene; Rb;
XX Caenorhabditis elegans; egi-43; retinoblastoma protein; cardiac muscle;
XX cell proliferation; cancer; cell proliferation; neural cell; paralysis;
XX neurodegenerative disease; Parkinson's disease; Huntington's disease;
XX Alzheimer's disease; motor neuron disorder; mouse; E1A; cell; cr2;
XX GTPase; SH3-binding motif.
XX
XX Rattus sp.
XX
XX US6323335-B1.
XX
XX 27-NOV-2001.
PD

```



```

XX PF 01-JUN-2000; 2000JS-0586472.
XX PF
XX PR 18-AUG-1995; 95US-0516859.
XX PR 17-MAR-2000; 2000JS-0528706.
XX PR 18-AUG-1994; 94US-0292683.
XX PR 06-MAR-1995; 95US-0399411.
XX PA (HUAN/) HUANG S.
XX FI Huang S;
XX DR WP1; 2002-096600/13.
XX PT Novel nucleic acid molecule comprises the code for a conserved domain
PT of retinoblastoma protein-interacting zinc finger protein, positive
PT regulatory domain i-binding factor involved in regulating gene
PT transcription
XX PS Example 3; Column 35; 93pp; English.
XX CC The invention relates to a nucleic acid molecule encoding a PR domain
CC peptide present in human and rat retinoblastoma protein-interacting zinc
CC finger protein (RIZ), PRDI-BF1 (human positive regulatory domain
CC i-binding factor 1), EVI-1 (human ecotropic viral integration site-1
CC myeloid transforming gene product) or Caenorhabditis elegans egl-43 gene
CC product. The DNA of the invention is involved in regulating transcription
CC of a target gene. RIZ acts as a cell-differentiation factor and modulates
CC a function of a cell by binding to retinoblastoma (Rb) protein, which is
CC involved in regulating cell proliferation in various human diseases of
CC conditions, e.g. cancer. RIZ regulates the growth of normal adult cardiac
CC muscle cells and prevents proliferation of surviving cells following
CC cardiac muscle cell death. RIZ is expressed in neural cells. This is
CC useful for healing after injury of neural tissue and for treating
CC neurodegenerative diseases such as Parkinson's, Huntington's or
CC Alzheimer's disease and paralysis, or motor neuron disorders through
CC induced or decreased proliferation of neural cells. The PR domain or RIZ
CC active fragment containing a PR domain is useful as a probe to identify
CC transcription factors or oncogenic proteins in a cell that bind the PR
CC domain. Sequences AAU0803-AAU0854 represent protein and peptide
CC molecules of the invention.
XX SQ Sequence 6 AA;

```

```

Query Match 55.3%; Score 21; DB 23; Length 6;
Best Local Similarity 80.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 4 DDLLE 8
DB 1 EDLJE 5

```

Search completed: October 20, 2003, 13:13:36  
Job time : 82 secs



```

; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; TITLE OF INVENTION: AND DETECTION OF CANCER
; FILE REFERENCE: 129.6USU1
; CURRENT APPLICATION NUMBER: US/09/834,765
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 349
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-349

```

```

Query Match      60.5%; Score 23; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 DLLEA 9
DB 5 DLLEA 9

RESULT 3
US-09-833-203-54
; Sequence 54, Application US/09833203
; Publication No. US20030166277A1
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; APPLICANT: Smith, Ernest S.
; TITLE OF INVENTION: Targeted Vaccine Delivery Systems
; FILE REFERENCE: 1821.0020001
; CURRENT APPLICATION NUMBER: US/09/833,203
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 60/196,472
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: misc feature
; OTHER INFORMATION: C35 peptides
US-09-833-203-54

```

```

Query Match      55.3%; Score 21; DB 12; Length 9;
Best Local Similarity 80.0%; Pred. No. 5.4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 DLLEA 9
DB 1 DLLEA 5

RESULT 4
US-10-166-225A-168
; Sequence 168, Application US/10166225A
; Publication No. US20030148416A1
; GENERAL INFORMATION:
; APPLICANT: BERRY, Alan
; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
US-09-833-203-54

```

```

; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: VELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/10/166,225A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 168
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Bradyrhizobium japonicum
US-10-166-225A-168

```

```

Query Match      52.6%; Score 20; DB 12; Length 5;
Best Local Similarity 60.0%; Pred. No. 5.4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 DDLLE 8
DB 1 DDLLE 5

```

```

RESULT 5
US-10-166-225A-169
; Sequence 169, Application US/10166225A
; Publication No. US20030148416A1
; GENERAL INFORMATION:
; APPLICANT: BERRY, Alan
; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: VELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/10/166,225A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 169
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Rhizobium sp. strain NGR234
US-10-166-225A-169

```

```

Query Match      52.6%; Score 20; DB 12; Length 5;
Best Local Similarity 60.0%; Pred. No. 5.4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 DDLLE 8
DB 1 DDLLE 5

```

```

RESULT 6
US-10-166-225A-170
; Sequence 170, Application US/10166225A
; Publication No. US20030148416A1
; GENERAL INFORMATION:
; APPLICANT: BERRY, Alan
; APPLICANT: BRETZEL, Werner
; APPLICANT: HUMBELIN, Markus
; APPLICANT: LOPEZ-ULIBARRI, Rual
; APPLICANT: MAYER, Anne F.
; APPLICANT: VELISEEV, Alexei A.
; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
; FILE REFERENCE: C38435/121966
; CURRENT APPLICATION NUMBER: US/10/166,225A
; CURRENT FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
US-10-166-225A-170

```

; SEQ ID NO 170  
 ; LENGTH: 5  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus stearothermophilus  
 US-10-166-225A-170

Query Match 52.6%; Score 20; DB 12; Length 5;  
 Best Local Similarity 60.0%; Pred. No. 5.4e+05;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
 DB 1 DDILD 5

RESULT 7  
 US-10-166-225A-171  
 ; Sequence 171, Application US/0166225A  
 ; Publication No. US20030148416A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BERRY, Alan  
 ; APPLICANT: BRETZEL, Werner  
 ; APPLICANT: HUMBELIN, Markus  
 ; APPLICANT: LOPEZ-ULIBARRI, Rual  
 ; APPLICANT: MAYER, Anne F.  
 ; APPLICANT: YELISEEV, Alexei A.  
 ; TITLE OF INVENTION: IMPROVED ISOPRENOL PRODUCTION  
 ; FILE REFERENCE: C38435/121966  
 ; CURRENT APPLICATION NUMBER: US/10/166,225A  
 ; CURRENT FILING DATE: 2002-06-05  
 ; NUMBER OF SEQ ID NOS: 197  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 171  
 ; LENGTH: 5  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus subtilis  
 US-10-166-225A-171

Query Match 52.6%; Score 20; DB 12; Length 5;  
 Best Local Similarity 60.0%; Pred. No. 5.4e+05;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
 DB 1 DDILD 5

RESULT 8  
 US-10-166-225A-172  
 ; Sequence 172, Application US/0166225A  
 ; Publication No. US20030148416A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BERRY, Alan  
 ; APPLICANT: BRETZEL, Werner  
 ; APPLICANT: HUMBELIN, Markus  
 ; APPLICANT: LOPEZ-ULIBARRI, Rual  
 ; APPLICANT: MAYER, Anne F.  
 ; APPLICANT: YELISEEV, Alexei A.  
 ; TITLE OF INVENTION: IMPROVED ISOPRENOL PRODUCTION  
 ; FILE REFERENCE: C38435/121966  
 ; CURRENT APPLICATION NUMBER: US/10/166,225A  
 ; CURRENT FILING DATE: 2002-06-05  
 ; NUMBER OF SEQ ID NOS: 197  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 172  
 ; LENGTH: 5  
 ; TYPE: PRT  
 ; ORGANISM: Escherichia coli  
 US-10-166-225A-172

Query Match 52.6%; Score 20; DB 12; Length 5;  
 Best Local Similarity 60.0%; Pred. No. 5.4e+05;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
 DB 1 DDILD 5

RESULT 9  
 US-10-166-225A-173  
 ; Sequence 173, Application US/0166225A  
 ; Publication No. US20030148416A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BERRY, Alan  
 ; APPLICANT: BRETZEL, Werner  
 ; APPLICANT: HUMBELIN, Markus  
 ; APPLICANT: LOPEZ-ULIBARRI, Rual  
 ; APPLICANT: MAYER, Anne F.  
 ; APPLICANT: YELISEEV, Alexei A.  
 ; TITLE OF INVENTION: IMPROVED ISOPRENOL PRODUCTION  
 ; FILE REFERENCE: C38435/121966  
 ; CURRENT APPLICATION NUMBER: US/10/166,225A  
 ; CURRENT FILING DATE: 2002-06-05  
 ; NUMBER OF SEQ ID NOS: 197  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 173  
 ; LENGTH: 5  
 ; TYPE: PRT  
 ; ORGANISM: Haemophilus influenzae  
 US-10-166-225A-173

Query Match 52.6%; Score 20; DB 12; Length 5;  
 Best Local Similarity 60.0%; Pred. No. 5.4e+05;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
 DB 1 DDILD 5

RESULT 10  
 US-10-271-708-8  
 ; Sequence 8, Application US/10271708  
 ; Publication No. US2003016220CA1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHOO, Yen  
 ; APPLICANT: LIU, Xiaohai  
 ; APPLICANT: SASASUBRAMANIAN, Shankar  
 ; APPLICANT: PATEL, Sachin D.  
 ; APPLICANT: ISALAN, Mark  
 ; TITLE OF INVENTION: METHODS FOR MODULATING TELOMERASE ACTIVITY  
 ; FILE REFERENCE: SABI-036/01US (8325-2014.01)  
 ; CURRENT APPLICATION NUMBER: US/10/271,708  
 ; CURRENT FILING DATE: 2002-10-15  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 8  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: F1 residue  
 US-10-271-708-8

Query Match 52.6%; Score 20; DB 12; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 5.4e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 DDLLE 8  
 DB 3 DDJTE 7

RESULT 11  
 US-09-801-784-19

Sequence 19, Application US/09801784  
Patent No. US20010014668A1  
GENERAL INFORMATION:  
APPLICANT: Cassels, Frederick J.  
Loomis-Price, Lawrence  
TITLE OF INVENTION: PEPTIDES FROM A CONSENSUS PEPTIDE OF  
E.  
COLI CS4-CFA/I FAMILY PROTEINS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hendricks and Assoc.  
STREET: P.O. Box 2509  
CITY: Fairfax  
STATE: VA  
COUNTRY: US  
ZIP: 22031  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/801,784  
FILING DATE: 09-Mar-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hendricks, Glenra M.  
REGISTRATION NUMBER: 32,535  
REFERENCE/DOCKET NUMBER: cas801  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 425-8405  
TELEFAX: (702) 425-8406  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: E. coli  
STRAIN: CS4-CFA/I  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-601-784-19  
Query Match 52.6%; Score 20; DB 9; Length 8;  
Best Local Similarity 80.0%; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0;  
Gaps 0;

QY 5 DLEA 9  
Db 4 DLEA 8

RESULT 12  
US-09-801-784-20  
Sequence 20, Application US/09801784  
Patent No. US20010014668A1  
GENERAL INFORMATION:  
APPLICANT: Cassels, Frederick J.  
Loomis-Price, Lawrence  
TITLE OF INVENTION: PEPTIDES FROM A CONSENSUS PEPTIDE OF  
E.  
COLI CS4-CFA/I FAMILY PROTEINS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hendricks and Assoc.  
STREET: P.O. Box 2509  
CITY: Fairfax  
STATE: VA  
COUNTRY: US  
ZIP: 22031  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/801,784  
FILING DATE: 09-Mar-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hendricks, Glenra M.  
REGISTRATION NUMBER: 32,535  
REFERENCE/DOCKET NUMBER: cas801  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 425-8405  
TELEFAX: (702) 425-8406  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: E. coli  
STRAIN: CS4-CFA/I  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-601-784-19

Sequence 20, Application US/09801784  
Patent No. US20010014668A1  
GENERAL INFORMATION:  
APPLICANT: Cassels, Frederick J.  
Loomis-Price, Lawrence  
TITLE OF INVENTION: PEPTIDES FROM A CONSENSUS PEPTIDE OF  
E.  
COLI CS4-CFA/I FAMILY PROTEINS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hendricks and Assoc.  
STREET: P.O. Box 2509  
CITY: Fairfax  
STATE: VA  
COUNTRY: US  
ZIP: 22031  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/801,784  
FILING DATE: 09-Mar-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hendricks, Glenra M.  
REGISTRATION NUMBER: 32,535  
REFERENCE/DOCKET NUMBER: cas801  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 425-8405  
TELEFAX: (702) 425-8406  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: E. coli  
STRAIN: CS4-CFA/I  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-801-784-20  
Query Match 52.6%; Score 20; DB 9; Length 8;  
Best Local Similarity 80.0%; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0;  
Gaps 0;

QY 5 DLEA 9  
Db 3 DLEA 7

RESULT 13  
US-09-801-784-21  
Sequence 21, Application US/09801784  
Patent No. US20010014668A1  
GENERAL INFORMATION:  
APPLICANT: Cassels, Frederick J.  
Loomis-Price, Lawrence  
TITLE OF INVENTION: PEPTIDES FROM A CONSENSUS PEPTIDE OF  
E.  
COLI CS4-CFA/I FAMILY PROTEINS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hendricks and Assoc.  
STREET: P.O. Box 2509  
CITY: Fairfax  
STATE: VA  
COUNTRY: US  
ZIP: 22031  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/801,784  
FILING DATE: 09-Mar-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hendricks, Glenra M.  
REGISTRATION NUMBER: 32,535  
REFERENCE/DOCKET NUMBER: cas801  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 425-8405  
TELEFAX: (702) 425-8406  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: E. coli  
STRAIN: CS4-CFA/I  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-801-784-21

TELEPHONE: (703) 425-8405  
TELEFAX: (702) 425-8406  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: E. coli  
STRAIN: CS4-CFA/1  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-801-784-21

Query Match 52.6%; Score 20; DB 9; Length 8;  
Best Local Similarity 80.0%; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLEA 9  
DB 2 DLLOA 6

RESULT 14  
US-09-801-784-22  
Sequence 22, Application US/99801784  
Patent No. US2001004668A  
GENERAL INFORMATION:  
APPLICANT: Cassels, Frederick J.  
J. Cassels-Price, Lawrence  
TITLE OF INVENTION: PEPTIDES FROM A CONSENSUS PEPTIDE OF  
E.  
COLI CS4-CFA/1 FAMILY PROTEINS

NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hendricks and Assoc.  
STREET: P.O. Box 2509  
CITY: Fairfax  
STATE: VA  
COUNTRY: US  
ZIP: 22031  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/801,784  
FILING DATE: 09-Mar-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hendricks, Glenn M.  
REGISTRATION NUMBER: 32,535  
REFERENCE/DOCKET NUMBER: cas801  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 425-8405  
TELEFAX: (702) 425-8406  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: E. coli  
STRAIN: CS4-CFA/1  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:

US-09-801-784-22

Query Match 52.6%; Score 20; DB 9; Length 8;  
Best Local Similarity 80.0%; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLEA 9  
DB 1 DLLOA 5

RESULT 15  
US-09-932-165-1239  
Sequence 1239, Application US/09932165  
Publication No. US20030134784A1  
GENERAL INFORMATION:  
APPLICANT: RATTANO, ARTHUR  
APPLICANT: CHALLITA-EID, PIA M.  
APPLICANT: FARIS, MARY  
APPLICANT: SAFFRAN, DOUGLAS  
APPLICANT: AFAR, DANIEL  
APPLICANT: HUBERT, RENE  
APPLICANT: LEVIN, ELANA  
APPLICANT: GE. WANGMAO  
APPLICANT: JAKOBOWITS, AVA  
TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED  
"93P2H3 AND CATF2E11 USEFUL IN TREATMENT AND  
DETECTION OF CANCER"  
FILE REFERENCE: 51:58-20014.00  
CURRENT APPLICATION NUMBER: US/09/932,165  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: 60/226,329  
PRIOR FILING DATE: 2000-08-17  
NUMBER OF SEQ ID NOS: 1508  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 1239  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Peptide motif  
US-09-932-165-1239

Query Match 52.6%; Score 20; DB 12; Length 9;  
Best Local Similarity 80.0%; Pred. No. 5.4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 DLLEA 9  
DB 4 DLLES 8

Search completed: October 20, 2003, 13:25:31  
Job time : 66 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 13:12:15 ; Search time 28 Seconds  
(without alignments)  
23.630 Million cell updates/sec

Title: US-09-489-760A-1

Perfect score: 38

Sequence: 1 VLKDD:LEA 3

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42312858 residues

Total number of hits satisfying chosen parameters: 77717

Minimum DB seq length: 0

Maximum DB seq length: 3

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- issued Patents AA:\*
- 1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/prodata/2/iaa/PCRTUS\_COMB.pep.\*
  - 6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	57.9	7	3	US-08-556-419-14
2	21	55.3	6	2	US-08-459-568-20
3	21	55.3	6	2	US-08-399-411-20
4	21	55.3	6	3	US-08-516-859A-20
5	21	55.3	6	4	US-09-566-472-20
6	21	55.3	6	4	US-09-528-706-20
7	21	55.3	8	2	US-08-459-568-75
8	21	55.3	8	2	US-08-399-411-75
9	21	55.3	8	3	US-08-516-859A-75
10	21	55.3	8	4	US-09-566-472-75
11	21	55.3	8	4	US-09-528-706-75
12	21	55.3	9	2	US-08-459-568-72
13	21	55.3	9	2	US-08-399-411-72
14	21	55.3	9	2	US-08-516-859A-72
15	21	55.3	9	3	US-09-566-472-72
16	21	55.3	9	3	US-09-528-706-72
17	21	55.3	9	3	US-08-516-859A-72
18	21	55.3	9	3	US-09-566-472-72
19	21	55.3	9	3	US-09-528-706-72
20	21	55.3	9	4	US-09-566-472-72
21	21	55.3	9	4	US-09-528-706-72
22	21	55.3	9	4	US-08-516-859A-72
23	21	55.3	9	4	US-09-566-472-72
24	20	52.6	6	3	US-09-085-819-11
25	20	52.6	6	4	US-09-808-126-11
26	20	52.6	6	4	US-09-803-951-11
27	20	52.6	8	3	US-08-747-599A-47

28	20	52.6	8	3	US-08-747-599A-50
29	20	52.6	8	3	US-08-747-599A-51
30	19	50.0	6	3	US-08-974-549A-373
31	19	50.0	6	4	US-08-912-951-140
32	19	50.0	7	4	US-09-187-859-1186
33	19	50.0	7	4	US-09-839-542B-1186
34	19	50.0	8	4	US-09-187-859-1189
35	19	50.0	8	4	US-09-187-859-2550
36	19	50.0	8	4	US-09-183-266A-35
37	19	50.0	8	4	US-09-183-266A-36
38	19	50.0	8	4	US-09-183-266A-38
39	19	50.0	8	4	US-09-839-542B-1189
40	19	50.0	8	4	US-09-839-542B-2550
41	19	50.0	9	1	US-07-663-413-15
42	19	50.0	9	1	US-08-059-530-15
43	19	50.0	9	1	US-08-214-650-49
44	19	50.0	9	1	US-08-467-083-20
45	19	50.0	9	1	US-08-414-417B-20

ALIGNMENTS

RESULT 1  
US-08-556-419-14  
; Sequence 14, Application US/08556419C  
; Patent No. 6093549  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Christopher  
; APPLICANT: Li, Xiao-Jiang  
; APPLICANT: Li, Shi-Hua  
; APPLICANT: Sharp, Alan  
; APPLICANT: Latham, Anthony  
; APPLICANT: Worley, Paul  
; APPLICANT: Snyder, Solomon  
; TITLE OF INVENTION: Huntingtin-associated protein  
; FILE REFERENCE: 01:07-52271  
; CURRENT APPLICATION NUMBER: US/08/556,419C  
; CURRENT FILING DATE: 1995-11-09  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-08-556-419-14

Query Match 57.9%; Score 22; DB 3; Length 7;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DLLL 8  
DB 1 DLLL 5

RESULT 2  
US-08-459-568-20  
; Sequence 20, Application US/08459568  
; Patent No. 5811304  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Shi  
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
; TITLE OF INVENTION: Zinc Finger Proteins  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,568  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/399,411  
FILING DATE: 06-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-459-568-20

Query Match 55.3%; Score 21; DB 2; Length 6;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8  
Db 1 EDLE 5

RESULT 3  
US-08-399-411-20  
Sequence 20, Application US/08399411  
Patent No. 5831008  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/399,411  
FILING DATE: 06-MAR-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-399-411-20

Query Match 55.3%; Score 21; DB 2; Length 6;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8  
Db 1 EDLE 5

RESULT 4  
US-08-516-859A-20  
Sequence 20, Application US/08516859A  
Patent No. 6069231  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/516,859A  
FILING DATE: 18-AUG-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/399,411  
FILING DATE: 06-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/292,683  
FILING DATE: 18-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1776  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-516-859A-20

Query Match 55.3%; Score 21; DB 3; Length 6;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8  
Db 1 EDLE 5

RESULT 5  
US-09-586-472-20  
Sequence 20, Application US/09586472  
Patent No. 6323335  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
Zinc Finger Proteins  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:



ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DCS  
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/528,706  
FILING DATE: 01-Jun-2000  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/528,706  
FILING DATE: 17-MAR-2000  
APPLICATION NUMBER: US 08/516,859  
FILING DATE: 18-AUG-1995  
APPLICATION NUMBER: US 08/399,411  
FILING DATE: 06-MAR-1995  
APPLICATION NUMBER: US 08/292,683  
FILING DATE: 18-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 4130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-586-472-20

Query Match 55.3%; Score 21; DB 4; Length 6;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8  
Db :|||||  
: EDLE 5

RESULT 6  
US-09-528-706-20  
Sequence 20, Application US/09528706  
Patent No. 6468985

GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DCS  
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/528,706  
FILING DATE:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/516,859  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/292,683  
FILING DATE: 18-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1776  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-09-528-706-20

Query Match 55.3%; Score 21; DB 4; Length 6;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8  
Db :|||||  
: EDLE 5

RESULT 7  
US-08-459-568-75  
Sequence 75, Application: US/08459568  
Patent No. 5811304

GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DCS  
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,568  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/399,411  
FILING DATE: 06-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-08-459-568-75

Query Match 55.3%; Score 21; DB 2; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 DDLE 8
   :||||
Db 1 EDLE 5

RESULT 8
US-08-399-411-75
; Sequence 75, Application US/08399411
; Patent No. 5831008
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent-in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,411
; FILING DATE: 06-MAR-1995
; CLASSIFICATION: 510
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-9949
; INFORMATION FOR SEQ ID NO: 75:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-399-411-75

```

```

Query Match 55.3%; Score 21; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8
   :||||
Db 1 EDLE 5

US-08-399-411-75

```

```

Query Match 55.3%; Score 21; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8
   :||||
Db 1 EDLE 5

RESULT 9
US-08-516-859A-75
; Sequence 75, Application US/08516859A
; Patent No. 6069231
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent-in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/516,859A
; FILING DATE: 18-AUG-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,683
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 75:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-516-859A-75

Query Match 55.3%; Score 21; DB 3; Length 8;
Best Local Similarity 80.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8
   :||||
Db 1 EDLE 5

RESULT 10
US-09-586-472-75
; Sequence 75, Application US/09586472
; Patent No. 6121135
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent-in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/586,472
; FILING DATE: 01-JUN-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/528,706
; FILING DATE: 17-MAR-2000
; APPLICATION NUMBER: US 08/516,859
; FILING DATE: 18-AUG-1995
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; APPLICATION NUMBER: US 08/292,683
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815

```

4  
REFERENCE/DOCKET NUMBER: P-LJ 4130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 75:  
US-09-596-472-75

Query Match 55.3%; Score 21; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
DB 1 EDLLE 5

RESULT 11  
US-09-528-706-75  
Sequence 75, Application US/39528706  
Patent No. 6468985  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/528.706  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/516.859  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/292.683  
FILING DATE: 18-JUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1776  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-528-706-75

Query Match 55.3%; Score 21; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
DB 1 EDLLE 5

RESULT 12  
US-08-459-568-72  
Sequence 72, Application US/08459568  
Patent No. 581304  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459.568  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/399.411  
FILING DATE: 06-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 72:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-459-568-72

Query Match 55.3%; Score 21; DB 2; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLE 8  
DB 1 EDLLE 5

RESULT 13  
US-08-459-568-78  
Sequence 78, Application US/08459568  
Patent No. 581304  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
US-08-459-568-72

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,568  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/399,411  
FILING DATE: 06-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-459-568-78

Query Match 55.3%; Score 21; DB 2; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8  
DB 1 EDLE 5

RESULT 14  
US-08-399-411-72  
Sequence 72, Application US/08399411  
Patent No. 5831008  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/399,411  
FILING DATE: 06-MAR-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 72:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-399-411-72

Query Match 55.3%; Score 21; DB 2; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8  
DB 1 EDLE 5

RESULT 15  
US-08-399-411-78  
Sequence 78, Application US/08399411  
Patent No. 5831008  
GENERAL INFORMATION:  
APPLICANT: Huang, Shi  
TITLE OF INVENTION: Retinoblastoma Protein - Interacting  
TITLE OF INVENTION: Zinc Finger Proteins  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/399,411  
FILING DATE: 06-MAR-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-399-411-78

Query Match 55.3%; Score 21; DB 2; Length 9;  
Best Local Similarity 80.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLE 8  
DB 1 EDLE 5

Search completed: October 20, 2003, 13:17:07  
Job time : 29 secs

```
; Entered [jdelaval 20-Oct-03 11:06]
SEQ1A
virdl.eal
```

```
; Entered [jdelaval 20-Oct-03 11:07]
SEQ1B
virdl.eal
```